MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

②

NAVAL POSTGRADUATE SCHOOL
Monterey, California

DTIC
S ELECTE D
DEC 1 6 1987

THESIS

FREQUENCY-SAMPLING DESIGN OF
TWO-DIMENSIONAL FIR DIGITAL FILTERS
WITH NONUNIFORM SAMPLES

by

William J. Rozwod

September 1987

Thesis Advisor: Charles W. Therrien

12   0   058

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION	1b RESTRICTIVE MARKINGS
UNCLASSIFIED	

2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT
	Approved for public release;
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	distribution is unlimited

4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)

6a NAME OF PERFORMING ORGANIZATION	6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION
Naval Postgraduate School	Code 62	Naval Postgraduate School

6c ADDRESS (City, State, and ZIP Code)	7b ADDRESS (City, State, and ZIP Code)
Monterey, California 93943-5000	Monterey, California 93943-5000

8a NAME OF FUNDING/SPONSORING ORGANIZATION	8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c ADDRESS (City, State and ZIP Code)	10 SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)
FREQUENCY-SAMPLING DESIGN OF TWO-DIMENSIONAL FIR DIGITAL FILTERS WITH NONUNIFORM SAMPLES

12 PERSONAL AUTHOR(S)
Rozwod, William J.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year Month Day)	15 PAGE COUNT
Engineer's Thesis	FROM _____ TO _____	1987 September	82

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB GROUP	Filters; Digital; Two-Dimensional; Finite
			Impulse Reponse (FIR); Frequency-Sampling

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Various approaches to the frequency-sampling design of two-dimensional FIR filters are analyzed. The IDFT approach requiring uniform sampling on a Cartesian grid is first described. A method which allows arbitrary placement of frequency samples but which does not satisfy the Haar condition is presented. Finally, a novel, computationally efficient method which allows nonuniform sampling and which always provides a unique design solution is presented. The new approach is compared with the other methods in terms of design flexibility, computational efficiency, and performance.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclassified

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
Prof. C.W. Therrien	(408) 646-3347	Code 62Ti

DD FORM 1473, 84 MAR
83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

Frequency-Sampling Design of Two-Dimensional
FIR Digital Filters With Nonuniform Samples

by

William J. Rozwod
Lieutenant, United States Navy
B.S., Rensselaer Polytechnic Institute, 1981

Submitted in partial fulfillment of the
requirements for the degrees of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING
and
ELECTRICAL ENGINEER

from the

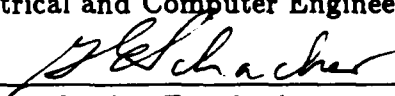NAVAL POSTGRADUATE SCHOOL
September 1987

Author: _____
William J. Rozwod

Approved by: _____
Charles W. Therrien, Thesis Advisor

_____
Roberto Cristi, Second Reader

_____
John R. Powers, Chairman, Department of
Electrical and Computer Engineering

_____
Gordon E. Schacher,
Dean of Science and Engineering

2

ABSTRACT

Various approaches to the frequency-sampling design of two-dimensional FIR filters are analyzed. The IDFT approach requiring uniform sampling on a Cartesian grid is first described. A method which allows arbitrary placement of frequency samples but which does not satisfy the Haar condition is presented. Finally, a novel, computationally efficient method which allows nonuniform sampling and which always provides a unique design solution is presented. The new approach is compared with the other methods in terms of design flexibility, computational efficiency, and performance.

Accesion For	
NTIS CRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

3

TABLE OF CONTENTS

5

ACKNOWLEDGEMENT

The author gratefully acknowledges the contributions and assistance provided by Prof. J.S. Lim, Massachusetts Institute of Technology, and Prof. C.W. Therrien, Naval Postgraduate School, during the course of this thesis research.

I. INTRODUCTION

Finite impulse response (FIR) filters have seen widespread use in the field of two-dimensional digital signal processing. Reasons for this include the inherent stability of FIR filters, the ability to achieve a linear (or zero) phase characteristic in the frequency response, and relative ease of design. Traditional approaches to the design of such filters include the window method, in which a smoothing window is applied to the Fourier series coefficients of the ideal frequency response; the frequency transformation method, in which a one-dimensional (1-D) prototype filter response is mapped into a function of two frequency variables using an appropriate transformation function; and the frequency-sampling method, in which desired frequency response values are specified at certain sample points in the frequency domain. Of the three methods described above, the frequency-sampling method is probably the least utilized and the least understood, and hence warrants further investigation. This thesis deals with a comparative analysis of various methods of designing two-dimensional (2-D) FIR filters using frequency-sampling techniques. In the analysis, some new methods for frequency-sampling design are presented.

The 2-D frequency-sampling design method is, in essence, an interpolation problem in two variables where the two variables are the two frequencies of interest, denoted as ω_1 and ω_2. The frequency response is specified at selected sample points in the (ω_1, ω_2) plane, and is expressed in terms of linearly independent basis functions. The solution to this interpolation problem involves finding the coefficients associated with the respective basis functions in order to meet the interpolating conditions. These coefficients are, in general, closely related to the impulse response. The various approaches to the frequency-sampling design

7

problem involve differing constraints placed on sample location in the (ω_1, ω_2) plane and differing sets of basis functions. Because the resulting frequency response takes on exactly the prescribed value at each frequency sample location, the response is somewhat predictable at the onset of the design process. A disadvantage of 2–D frequency–sampling design, however, is that the existence and uniqueness of a solution is dependent upon the location of selected frequency samples and the basis functions specified; in other words, degenerate cases may arise in which no solution exists to the bivariate interpolation problem.

FIR filter design techniques denoted as "frequency–sampling" traditionally involve sampling a desired frequency response at equally-spaced intervals and then solving for the impulse response coefficients using an inverse discrete Fourier transform (IDFT) [Ref. 1]. For the 2–D case, the process is similar, with the desired frequency response sampled at equally spaced points on a Cartesian grid in the (ω_1, ω_2) plane [Refs. 2, 3]. This approach will be referred to as the "uniform sampling" method in this thesis. There are two significant benefits to such an approach in two dimensions. First, with frequency samples placed at the vertices of an $N \times N$ Cartesian grid, the existence of a unique filter satisfying the interpolating conditions and possessing an impulse response with an $N \times N$ region of support is guaranteed. Second, use of the IDFT, or alternatively, the fast Fourier transform (FFT) algorithm, assures a computationally efficient design process. The principal disadvantage of the approach is that there is no flexibility in the placement of frequency samples. The approach is presented in more detail in Chapter II.

A different approach to frequency–sampling design of 2–D FIR filters involves allowing frequency samples to occur anywhere in the (ω_1, ω_2) plane and fixing the region of support for the filter impulse response. This approach is presented in Chapter III and will be denoted as the "arbitrary sampling" technique. The

8

method is quite flexible since samples can be arranged to enhance certain desired features in the resulting frequency response, such as flat passband, sharp roll-off, equi-ripple behavior, etc. The design, however, involves the solution of a large system of linear equations for a filter of any practical order and hence it is quite computationally intensive. Additionally, with no constraints on the location of samples, degenerate cases may arise in which no solution exists which meets all interpolating conditions for the specified impulse response region of support.

A novel 2–D frequency–sampling design technique is presented in Chapter IV, and is referred to as the "nonuniform sampling" method. This method places certain constraints on frequency sample location but still possesses much greater flexibility than the uniform sampling approach. Rather than solving one large system of equations, this technique involves the solution of several smaller systems of linear equations. Given the constraints on the sampling geometry, existence and uniqueness of a solution is guaranteed. Additionally, the design method is less computationally intensive than the arbitrary sampling case but more intensive, in general, than the uniform sampling case.

In comparing and contrasting these various approaches to 2–D frequency–sampling design, several interesting points arise. First, frequency-sampling design techniques involve some sort of decision–making process in determining where to place the samples in order to best match the desired frequency response. Methods which are more constrained, such as the uniform sampling approach, involve a lesser degree of decision–making and hence can be considered easier to use from that perspective. Second, by reducing the number of constraints on the placement of frequency samples, flexibility in matching the desired frequency response is gained. Finally, and perhaps most significantly, computational efficiency in determining filter parameters can be directly related to the number of constraints on sample

9

location. Of the techniques discussed here, those which are more constrained tend to be more computationally efficient. Thus there are tradeoffs to be considered when selecting one of these techniques. In this thesis, a detailed analysis of these concepts is performed and the quality of these frequency-sampling methods in matching desired filter characteristics is investigated through design examples.

II. DESIGNS WITH UNIFORM SAMPLING ON A CARTESIAN GRID

The approach to frequency-sampling design of 2-D FIR filters discussed in this chapter is the one referred to as "frequency-sampling" in various texts [Refs. 1, 2]. It will be referred to the "uniform sampling" method through the remainder of this thesis. This method has significant computational advantages through use of an inverse discrete Fourier transform (IDFT) approach. Additionally, this method will never result in any degenerate cases, i.e. situations in which a filter satisfying each of the interpolating conditions cannot be found. However, due to constraints on the location of samples, the method is rather inflexible.

A. APPROACH

The transfer function of a two-dimensional FIR filter with impulse response support in the first quadrant can be defined as

$$H(z_1, z_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} h(n_1, n_2) z_1^{-n_1} z_2^{-n_2}. \qquad (2.1)$$

The frequency response of a such a filter is obtained by evaluating the transfer function on the unit bicircle through the substitutions $z_1 = e^{j\omega_1}$ and $z_2 = e^{j\omega_2}$, yielding

$$H(\omega_1, \omega_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} h(n_1, n_2) e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}. \qquad (2.2)$$

The frequency-sampling filter design problem involves determination of a sequence $h(n_1, n_2)$ which will result in a filter matching a desired frequency response $H_d(\omega_1, \omega_2)$ at specified sample points in the (ω_1, ω_2) plane.

The initial step in the design process is to choose the filter order, and hence the parameters N_1 and N_2. Often it is desirable to specify $N_1 = N_2 = N$ to

11

ensure a result with equal dependence on the frequencies ω_1 and ω_2. The desired frequency response $H_d(\omega_1, \omega_2)$ is sampled at the vertices of a uniformly spaced $N_1 \times N_2$ Cartesian grid on the region $\{0 \le \omega_1 \le 2\pi; \ 0 \le \omega_2 \le 2\pi\}$ of the (ω_1, ω_2) plane. The resulting frequency samples can be expressed in 2-D sequence form as

$$\tilde{H}(k_1, k_2) = H_d(\omega_1, \omega_2)\Big|_{\omega_1 = \frac{2\pi k_1}{N_1}, \omega_2 = \frac{2\pi k_2}{N_2}},$$

$$k_1 = 0, 1, \ldots, N_1 - 1 \quad \text{and} \quad k_2 = 0, 1, \ldots, N_2 - 1. \tag{2.3}$$

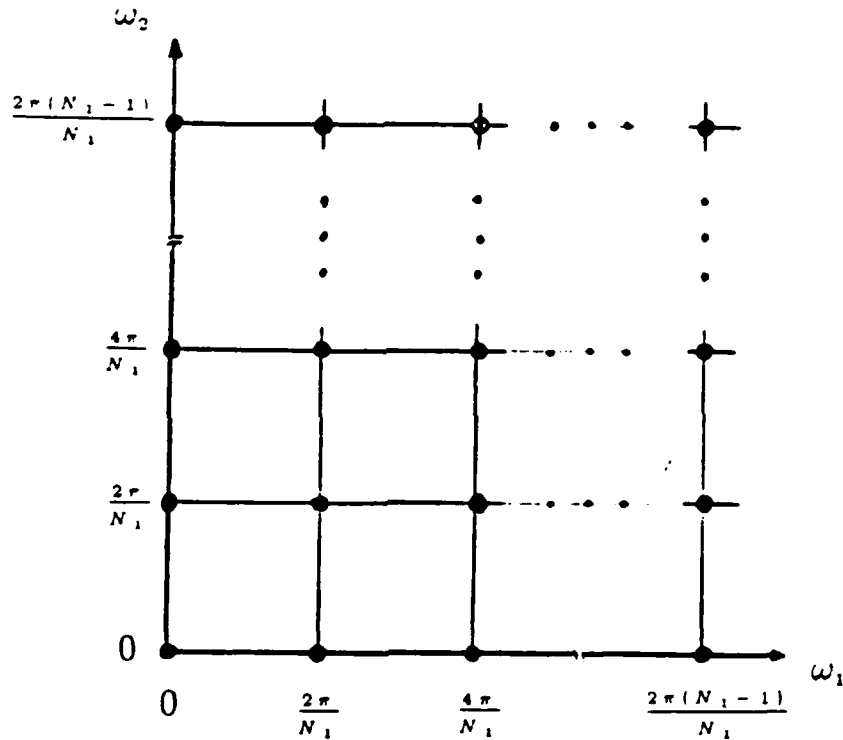The resulting sampling geometry is illustrated in Figure 2.1.

Figure 2.1. Arrangement of Samples for Uniform Sampling Method.

Because an IDFT approach is to be used in computing $h(n_1, n_2)$, and since the IDFT is defined for first-quadrant sequences only, the desired frequency response should possess a linear phase characteristic such that

$$H_d(\omega_1, \omega_2) = |H_d(\omega_1, \omega_2)| e^{-j\omega_1 M_1} e^{-j\omega_2 M_2}, \tag{2.4}$$

12

where

$$M_1 = \begin{cases} N_1/2 & N_1 \text{ even} \\ (N_1 - 1)/2 & N_1 \text{ odd} \end{cases} \tag{2.5}$$

and

$$M_2 = \begin{cases} N_2/2 & N_2 \text{ even} \\ (N_2 - 1)/2 & N_2 \text{ odd} \end{cases} . \tag{2.6}$$

Additionally, $H_d(\omega_1, \omega_2)$ must possess the symmetry conditions

$$\begin{aligned} |H(k_1, k_2) &= |H(k_1, N_2 - k_2)| \\ &= |H(N_1 - k_1, k_2)| \end{aligned} \tag{2.7}$$

in order to ensure a purely real impulse response.

The next step is to compute $h(n_1, n_2)$ by taking the IDFT of the sampled desired frequency response:

$$\begin{aligned} h(n_1, n_2) &= \text{IDFT}[\tilde{H}(k_1, k_2)] \\ &= \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \tilde{H}(k_1, k_2) W_{N_1}^{-n_1 k_1} W_{N_2}^{-n_2 k_2}, \end{aligned} \tag{2.8}$$

where

$$W_{N_1} = e^{-j\frac{2\pi}{N_1}} \quad \text{and} \quad W_{N_2} = e^{-j\frac{2\pi}{N_2}}.$$

The resulting sequence $h(n_1, n_2)$ is purely real and represents the impulse response of the designed filter. Such a filter possesses a linear phase characteristic. A zero-phase characteristic can be obtained by shifting the impulse response so that it is centered about the origin such that

$$h'(m_1, m_2) = h(m_1 + M_1, m_2 + M_2), \tag{2.9}$$

where $h'(m_1, m_2)$ represents the zero-phase impulse response.

13

To see the relationship between the resulting filter frequency response and the specified sample sequence $\tilde{H}(k_1, k_2)$, Eq. (2.8) is substituted into Eq. (2.2):

$$H(\omega_1, \omega_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \left[\frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \tilde{H}(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2} \right] e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}$$

$$= \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \tilde{H}(k_1, k_2) \sum_{n_1=0}^{N_1-1} (W_{N_1}^{-k_1} e^{-j\omega_1})^{n_1} \sum_{n_2=0}^{N_2-1} (W_{N_2}^{-k_2} e^{-j\omega_2})^{n_2}$$

$$= \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \frac{(1 - e^{-j\omega_1 N_1})(1 - e^{-j\omega_2 N_2})}{N_1 N_2} \frac{\tilde{H}(k_1, k_2)}{(1 - W_{N_1}^{-k_1} e^{-j\omega_1})(1 - W_{N_2}^{-k_2} e^{-j\omega_2})}.$$

$$(2.10)$$

The above expression can be rewritten in the form

$$H(\omega_1, \omega_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \tilde{H}(k_1, k_2) \Phi(\omega_1 - \frac{2\pi}{N_1} k_1, \omega_2 - \frac{2\pi}{N_2} k_2), \qquad (2.11)$$

where

$$\Phi(\omega_1, \omega_2) = \frac{(1 - e^{-j\omega_1 N_1})(1 - e^{-j\omega_2 N_2})}{N_1 N_2 (1 - e^{-j\omega_1})(1 - e^{-j\omega_2})}$$
$$\cdot e^{-j\omega_1 (\frac{N_1-1}{2})} e^{-j\omega_2 (\frac{N_2-1}{2})} e^{j\pi k_1 (1 - \frac{1}{N_1})} e^{j\pi k_2 (1 - \frac{1}{N_2})}$$

which simplifies to

$$\Phi(\omega_1, \omega_2) = \frac{\sin(\omega_1 N_1 / 2) \sin(\omega_2 N_2 / 2)}{N_1 N_2 \sin(\omega_1 / 2) \sin(\omega_2 / 2)}$$
$$\cdot e^{-j\omega_1 (\frac{N_1-1}{2})} e^{-j\omega_2 (\frac{N_2-1}{2})} e^{j\pi k_1 (1 - \frac{1}{N_1})} e^{j\pi k_2 (1 - \frac{1}{N_2})} \qquad (2.12)$$

for the case where N_1 and N_2 are both odd integers. This analysis is a 2-D extension of operations performed for the one-dimensional case [Ref. 1: pp. 97-98]. Since

$$\Phi(\frac{2\pi}{N_1} k_1, \frac{2\pi}{N_2} k_2) = \begin{cases} 1 & (k_1, k_2) = (0, 0) \\ 0 & k_1 = 1, 2, \ldots, N_1 - 1 \quad \text{and} \quad k_2 = 1, 2, \ldots, N_2 - 1 \end{cases},$$

the frequency response is matched exactly at the original sample points. That is,

$$H(\omega_1, \omega_2)\big|_{\omega_1 = \frac{2\pi k_1}{N_1}, \omega_2 = \frac{2\pi k_2}{N_2}} = \tilde{H}(k_1, k_2), \qquad (2.13)$$

14

$$k_1 = 0, 1, \ldots, N_1 - 1 \quad \text{and} \quad k_2 = 0, 1, \ldots, N_2 - 1.$$

Thus, the design process entails a bivariate interpolation in the variables ω_1 and ω_2, and, due to use of the IDFT algorithm, the interpolating functions are of the form of Eq. (2.12).

In using this filter design method, the choice of the desired frequency response $H_d(\omega_1, \omega_2)$ has significant effects on the quality of the resulting filter. For instance, if an "ideal" lowpass filter characteristic with the circular symmetry shown in Figure 2.2 is used for $H_d(\omega_1, \omega_2)$, the passband and stopband of the resulting filter frequency response exhibit substantial ripple.
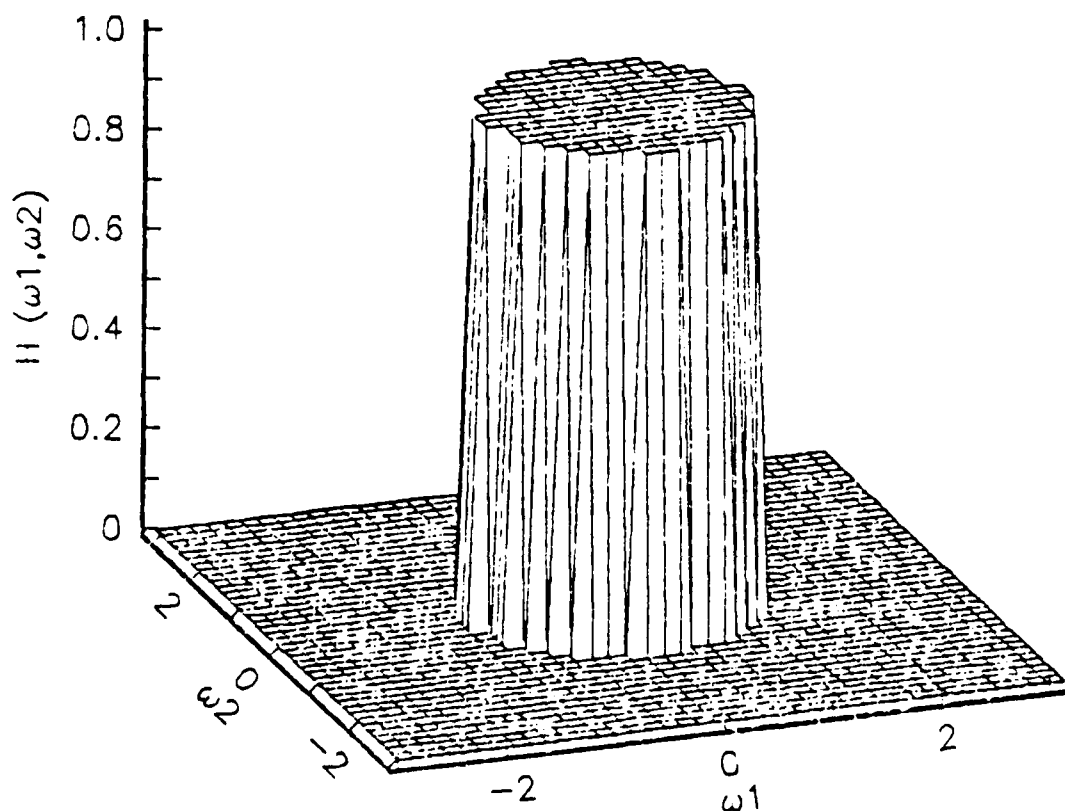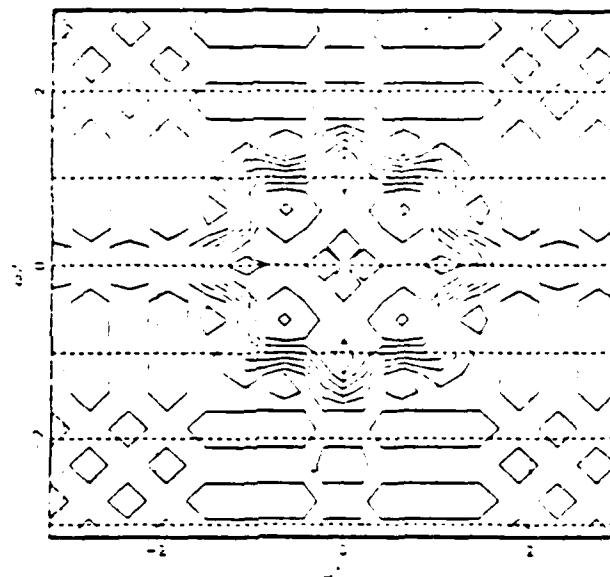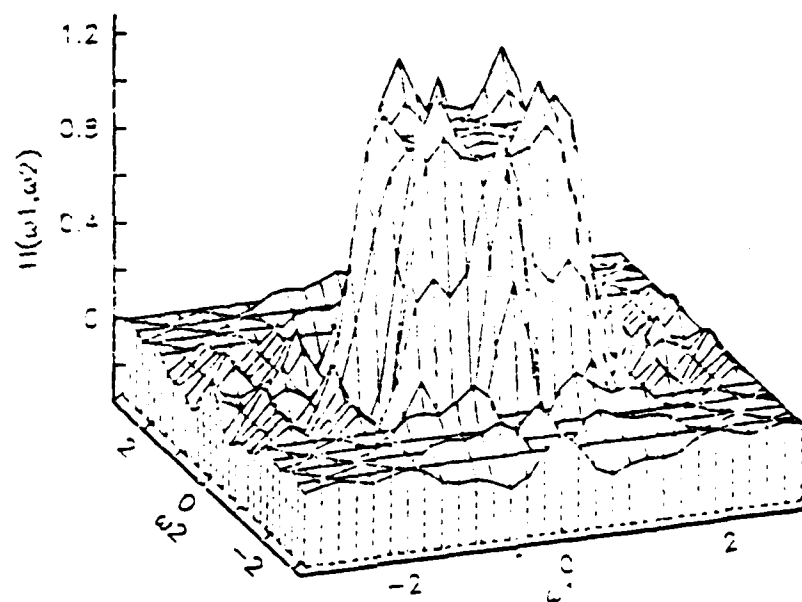

Figure 2.2. Ideal Lowpass Filter Frequency Response.

15

Figure 2.3 illustrates contour and perspective plots of the frequency response arising from such a design. In this case, a 15 × 15 point filter was designed by uniformly sampling the desired frequency response of Figure 2.2.

The frequency response of the resulting filter can be improved considerably if a transition region is specified for $H_d(\omega_1, \omega_2)$. Increasing the width of such a transition region reduces passband and stopband ripple but does so at the expense of a less sharp roll-off characteristic in the resulting filter response. The values assigned to $\tilde{H}(k_1, k_2)$ for points (k_1, k_2) located in this transition region significantly influence the behavior of the resulting filter. The selection of these values in order to satisfy some specific optimization criterion for the frequency response, such as best equi-ripple approximation to an ideal response, is not trivial. Since $H(\omega_1, \omega_2)$ is a linear function of the values $\tilde{H}(k_1, k_2)$, as seen in Eq. (2.10), linear optimization techniques can be employed in selecting the transition region values of $\tilde{H}(k_1, k_2)$. Such techniques for the one-dimensional case are rather involved [Ref. 4]; for the 2-D case, the problem is much more computationally intensive [Ref. 2], and hence, has not been addressed. Specifying $H_d(\omega_1, \omega_2)$ to have a smooth, linear transition between the passband and stopband regions is the simplest thing to do and generally works well. The magnitude characteristic of a desired frequency response with a linear, circularly symmetric transition region is shown in Figure 2.4. A filter can then be designed by sampling this response on the uniform Cartesian grid specified in Eq. (2.3). Contour and perspective plots of a 15 × 15 filter designed in this manner are shown in Figure 2.5. In comparing this result to that of Figure 2.3, it is evident that inclusion of a transition region greatly reduces passband and stopband ripple.

16

(a)

(b)

Figure 2.3. Frequency Response Arising From Sampling Ideal Lowpass
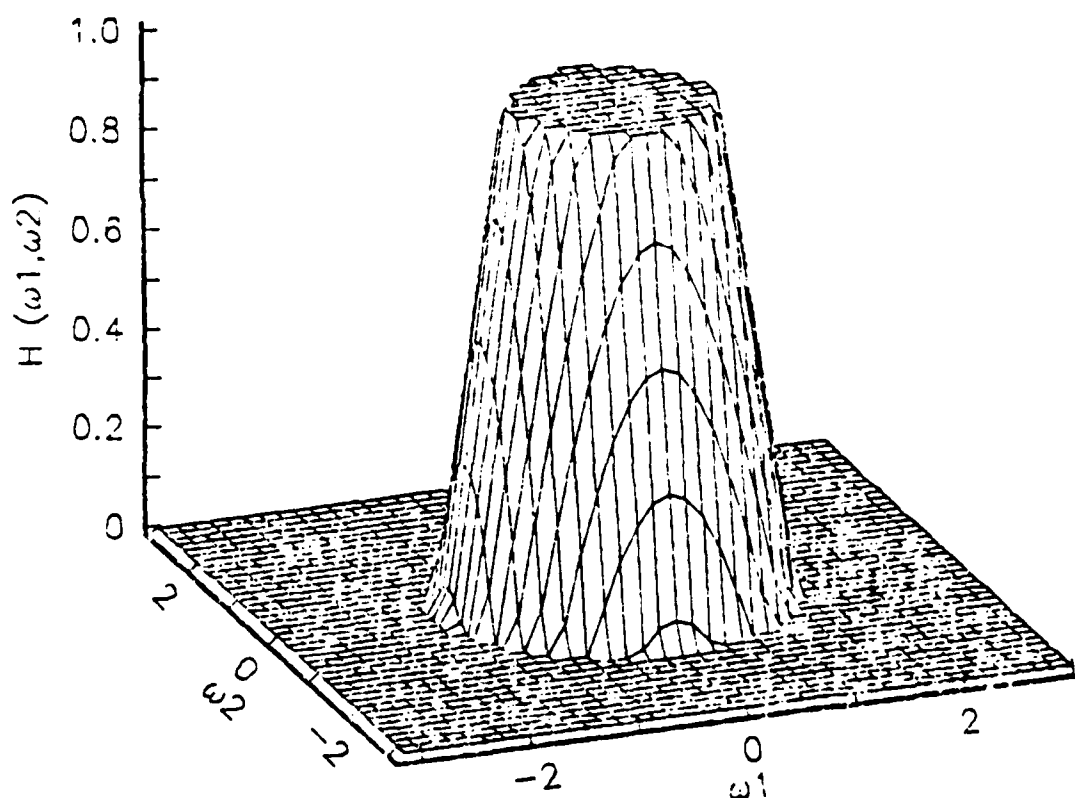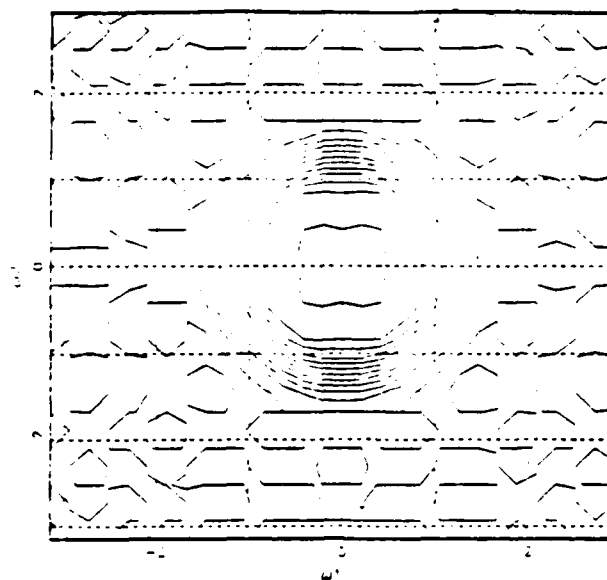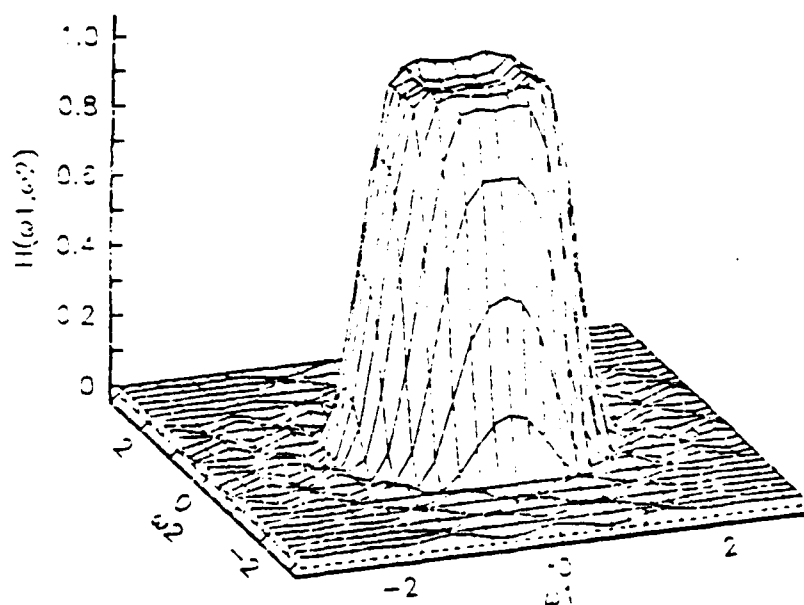Characteristic With No Transition Band.

17

Figure 2.4. Desired Lowpass Frequency Response Characteristic
With Linear Transition Band.

B.   EXISTENCE AND UNIQUENESS

A significant property of the uniform sampling design is that a unique filter satisfying each of the interpolating conditions specified in Eq. (2.3) is guaranteed to exist. In general, existence and uniqueness of such an interpolation result in two dimensions is not guaranteed; this can be a significant theoretical and practical problem. The point is discussed in further detail in Chapter III. A proof of existence of a unique $h(n_1, n_2)$ using the frequency-sampling method with uniformly spaced samples follows.

18

(a)

(b)

Figure 2.5. Frequency Response Arising From Sampling Desired Lowpass
Characteristic With Linear Slope Transition Band.

19

Equation (2.8) is used to determine $h(n_1, n_2)$. This can be rewritten as

$$h(n_1, n_2) = \frac{1}{N_1} \sum_{k_1=0}^{N_1-1} \left[\frac{1}{N_2} \sum_{k_2=0}^{N_2-1} \tilde{H}(k_1, k_2) W_{N_2}^{-k_2 n_2} \right] W_{N_1}^{-k_1 n_1}. \quad (2.14)$$

Now define

$$G(k_1, n_2) = \frac{1}{N_2} \sum_{k_2=0}^{N_2-1} \tilde{H}(k_1, k_2) W_{N_2}^{-k_2 n_2}. \quad (2.15)$$

Substituting Eq. (2.15) into Eq. (2.14), the following expression for the impulse response results:

$$h(n_1, n_2) = \frac{1}{N_1} \sum_{k_1=0}^{N_1-1} G(k_1, n_2) W_{N_1}^{-k_1 n_1}. \quad (2.16)$$

Thus the 2–D IDFT of $\tilde{H}(k_1, k_2)$ can be determined in a two-step process involving one-dimensional transforms known as row-column decomposition [Ref. 5: pp. 75-76].

The first step in the row-column decomposition process is to perform the solution of a total of N_1 systems of N_2 equations of the form of Eq. (2.15) using a 1–D IDFT algorithm. In matrix form, this can be written as

$$\begin{bmatrix} G(k_1, 0) \\ G(k_1, 1) \\ \vdots \\ G(k_1, N_2-1) \end{bmatrix} = \frac{1}{N_2} \begin{bmatrix} W_{N_2}^0 & W_{N_2}^0 & \cdots & W_{N_2}^0 \\ W_{N_2}^0 & W_{N_2}^{-1} & \cdots & W_{N_2}^{-(N_2-1)} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N_2}^0 & W_{N_2}^{-(N_2-1)} & \cdots & W_{N_2}^{-1} \end{bmatrix} \begin{bmatrix} \tilde{H}(k_1, 0) \\ \tilde{H}(k_1, 1) \\ \vdots \\ \tilde{H}(k_1, N_2-1) \end{bmatrix}, \quad (2.17)$$

which can be expressed in vector-matrix notation as

$$\mathbf{G} = \frac{1}{N_2} \mathbf{W}_2 \tilde{\mathbf{H}}. \quad (2.18)$$

It can be shown [Ref. 6: pp. 44-45] that a scaled version of $\mathbf{W}_2$, namely $\mathbf{W}_2/\sqrt{N_2}$, is unitary and hence non-singular. Therefore a unique solution for $\mathbf{G}$ exists for each k_i, $i = 0, 1, \ldots, N_1 - 1$.

20

The second step in the computation involves the solution of N_2 systems of N_1 equations of the form of Eq. (2.16). In matrix form, this becomes

$$\begin{bmatrix} h(0,n_2) \\ h(1,n_2) \\ \vdots \\ h(N_1-1,n_2) \end{bmatrix} = \frac{1}{N_1} \begin{bmatrix} W_{N_1}^0 & W_{N_1}^0 & \cdots & W_{N_1}^0 \\ W_{N_1}^0 & W_{N_1}^{-1} & \cdots & W_{N_1}^{-(N_1-1)} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N_1}^0 & W_{N_1}^{-(N_1-1)} & \cdots & W_{N_1}^{-1} \end{bmatrix} \begin{bmatrix} G(0,n_2) \\ G(1,n_2) \\ \vdots \\ G(N_1-1,n_2) \end{bmatrix},$$

(2.19)

which in vector-matrix form is

$$\mathbf{h} = \frac{1}{N_1}\mathbf{W}_1\mathbf{G}.$$

(2.20)

Since $\mathbf{G}$ was uniquely determined for each value of k_1, and since $\mathbf{W}_1$ is also non-singular, a unique solution for $\mathbf{h}$ exists for each n_j, $\quad j = 0, 1, \ldots, N_2 - 1$. Therefore, a unique solution for $h(n_1, n_2)$ is obtained from the sequence $\tilde{H}(k_1, k_2)$, which represents the samples of a desired frequency response at the vertices of a uniform Cartesian grid in the (ω_1, ω_2) plane.

C. COMPUTATIONS

While there is no flexibility in the choice of sample point locations using this IDFT approach, the method does offer some significant computational advantages. One criterion often used to measure computational efficiency in solving a problem is to determine the number of complex operations involved. While there exist many methods of computing the 2–D IDFT, the use of row-column decomposition and the IDFT directly (vice the FFT algorithm) is a reasonably efficient approach for the design of filters of practical order and is used in the following analysis.

The initial step in the row-column decomposition approach is to solve a total of N_1 systems of the form of Eq. (2.17). Solution of each system via the IDFT involves a total of N_2^2 multiplications and additions. Therefore, a total of $N_1 N_2^2$

floating point operations are required in this step, where a floating point operation is considered one multiplication and one addition.

The second part of the row-column decomposition approach involves the solution of N_2 systems of the form of Eq. (2.19). Solution of each system via the IDFT requires N_1^2 operations. Therefore, $N_1^2 N_2$ floating operations are involved in the second step.

The total number of floating point operations required in the solution for the impulse response coefficients of an $N_1 \times N_2$ frequency-sampling FIR filter design using uniform sampling on a Cartesian grid and a row-column decomposition, IDFT approach, is

$$C_{\text{uniform}} = N_1 N_2^2 + N_1^2 N_2. \tag{2.21}$$

For the case in which $N_1 = N_2 = N$, this reduces to

$$C_{\text{uniform}} = 2N^3. \tag{2.22}$$

In summary, the uniform sampling, IDFT approach to 2–D frequency sampling filter design has the benefit of computational efficiency, yet lacks flexibility in the choice of frequency samples. The next two chapters will present methods which achieve greater flexibility by allowing nonuniform spacing of samples.

III.   DESIGNS WITH ARBITRARILY PLACED SAMPLES

In Chapter II a frequency-sampling 2-D FIR filter design method was presented in which samples were placed on a uniform Cartesian grid in the (ω_1, ω_2) plane. Due to the constraints on sample placement, however, the method was not particularly flexible. If the desired frequency response $H_d(\omega_1, \omega_2)$ was fixed, no free parameters were available in optimizing the frequency response of the resulting filter. In this chapter, an alternate frequency-sampling design technique for 2-D FIR filters is presented in which samples are no longer constrained to a uniform Cartesian grid but can be placed anywhere in the (ω_1, ω_2) plane. This approach will be referred to as the "arbitrary-sampling" method. Allowing the samples to occur anywhere in the (ω_1, ω_2) plane introduces additional free parameters; hence, linear optimization techniques can be employed without varying $H_d(\omega_1, \omega_2)$, the desired frequency response to be sampled. Since the locations of the frequency samples are much less constrained, the method provides greater potential for achieving a desirable frequency response characteristic in the resulting filter. Since the placement of samples on a uniform Cartesian grid is a special case of the arbitrary-sampling method, the arbitrary-sampling method can achieve results at least as good as the uniform-sampling method for a fixed impulse response region of support and a fixed number of samples. However, the arbitrary-sampling method possesses a few significant drawbacks, such as computational complexity and the possibility of degenerate cases in which the interpolating conditions cannot be met. These must be considered in the selection of a method for filter design.

23

A. APPROACH

The arbitrary-sampling method can be used to produce 2-D FIR filters with a variety of impulse response regions of support. This region of support is one of the parameters to be specified at the outset of the design process. For illustration purposes, a filter with an $N_1 \times N_2$ rectangularly-shaped impulse response region of support centered on the origin of the (n_1, n_2) plane is specified, where N_1 and N_2 are odd integers. The transfer function of such a filter can be expressed as

$$H(z_1, z_2) = \sum_{n_1 = -M_1}^{M_1} \sum_{n_2 = -M_2}^{M_2} h(n_1, n_2) z_1^{-n_1} z_2^{-n_2}, \tag{3.1}$$

where $h(n_1, n_2)$ is the impulse response to be determined, $M_1 = (N_1 - 1)/2$ and $M_2 = (N_2 - 1)/2$. The parameters M_1 and M_2 will be used often in the remainder of this analysis, so the above relationships are worth noting. Evaluation of Eq. (3.1) on the unit bicircle yields the following expression for the frequency response:

$$H(\omega_1, \omega_2) = \sum_{n_1 = -M_1}^{M_1} \sum_{n_2 = -M_2}^{M_2} h(n_1, n_2) e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}. \tag{3.2}$$

To ensure a zero-phase characteristic in the resulting frequency response, $h(n_1, n_2)$ is specified to possess quadrant symmetry such that

$$h(n_1, n_2) = h(-n_1, n_2) = h(n_1, -n_2). \tag{3.3}$$

Application of Eq. (3.3) to Eq. (3.2) produces the following expression:

$$H(\omega_1, \omega_2) = h(0,0) + \sum_{n_1=1}^{M_1} 2h(n_1, 0) \cos \omega_1 n_1 +$$

$$\sum_{n_2=1}^{M_2} 2h(0, n_2) \cos \omega_2 n_2 + \sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} 4h(n_1, n_2) \cos \omega_1 n_1 \cos \omega_2 n_2. \tag{3.4}$$

24

This can be rewritten as

$$H(\omega_1, \omega_2) = \sum_{n_1=0}^{M_1} \sum_{n_2=0}^{M_2} a(n_1, n_2) \cos \omega_1 n_1 \cos \omega_2 n_2, \tag{3.5}$$

where

$$a(n_1, n_2) = \begin{cases} h(n_1, n_2) & n_1 = 0, n_2 = 0 \\ 2h(n_1, n_2) & n_1 \neq 0, n_2 = 0 \\ 2h(n_1, n_2) & n_1 = 0, n_2 \neq 0 \\ 4h(n_1, n_2) & n_1 \neq 0, n_2 \neq 0 \end{cases}. \tag{3.6}$$

The basis functions $\{\psi_i(\omega_1, \omega_2)\}$ are defined such that

$$\psi_i(\omega_1, \omega_2) = \cos \omega_1 n_1 \cos \omega_2 n_2, \tag{3.7}$$

where

$$i = n_1(M_2 + 1) + n_2 + 1; \qquad n_1 = 0, 1, \ldots, M_1;$$
$$n_2 = 0, 1, \ldots, M_2. \tag{3.8}$$

By arranging the 2-D sequence $a(n_1, n_2)$ into the one-dimensional sequence $a(i)$, the frequency response can then be expressed as

$$H(\omega_1, \omega_2) = \sum_{i=1}^{R} a(i) \psi_i(\omega_1, \omega_2), \tag{3.9}$$

where

$$R = (M_1 + 1)(M_2 + 1). \tag{3.10}$$

From Eq. (3.9), the elements of the sequence $a(i)$ represent a total of R free parameters to be determined. Therefore, the frequency response is to be specified at a total of R distinct sample points in the (ω_1, ω_2) plane.

The initial step in the design process is to specify $H(\omega_1, \omega_2)$ at R distinct sample points in the region $K = \{(\omega_1, \omega_2) : 0 \leq \omega_1 \leq \pi; 0 \leq \omega_2 \leq \pi\}$. An example of such a sampling geometry is illustrated in Figure 3.1.

No constraints other than the distinctness requirement are placed on the location of samples within this region. The form of Eq. (3.5) implies that the frequency
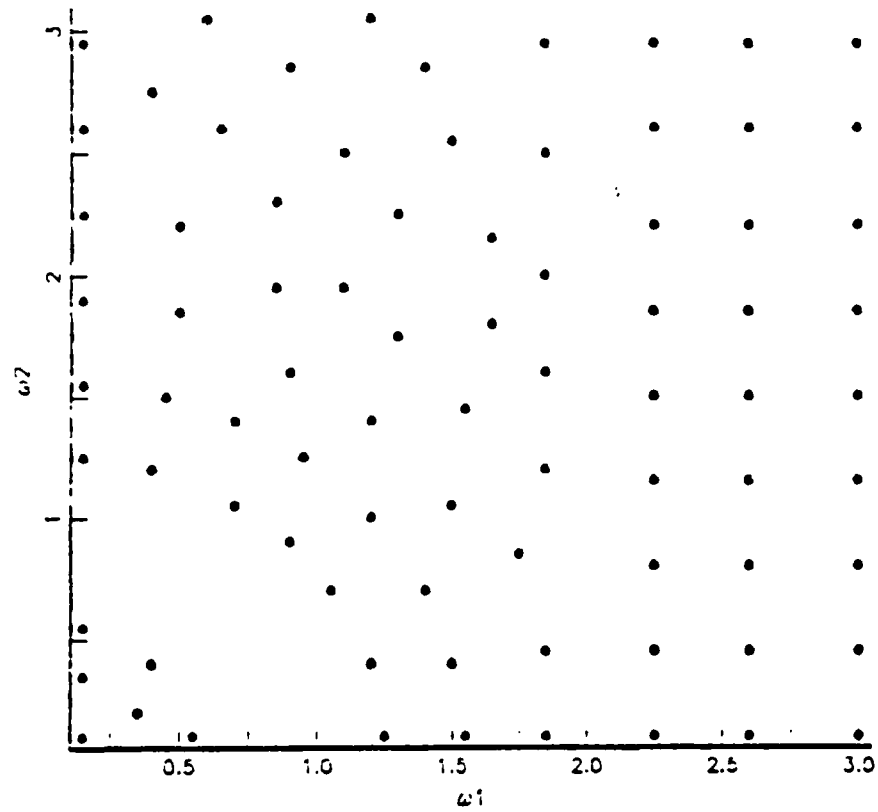
25

Figure 3.1. Possible Arrangement of Samples for Arbitrary Sampling Design.

response possesses quadrantal symmetry and periodicity of 2π in both frequencies. Hence, $H(\omega_1, \omega_2)$ is defined for all (ω_1, ω_2).

The next step in the design process is to solve for the sequence $a(n_1, n_2)$. Application of Eq. (3.9) at each specified sample location $\{(\omega_1, \omega_2)_i \; ; \quad i = 1, 2, \ldots, R\}$ results in the following system of linear equations:

$$
\begin{bmatrix} H(\omega_1, \omega_2)_1 \\ H(\omega_1, \omega_2)_2 \\ \vdots \\ H(\omega_1, \omega_2)_R \end{bmatrix} = \begin{bmatrix} \psi_1(\omega_1, \omega_2)_1 & \psi_2(\omega_1, \omega_2)_1 & \ldots & \psi_R(\omega_1, \omega_2)_1 \\ \psi_1(\omega_1, \omega_2)_2 & \psi_2(\omega_1, \omega_2)_2 & \ldots & \psi_R(\omega_1, \omega_2)_2 \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\omega_1, \omega_2)_R & \psi_2(\omega_1, \omega_2)_R & \ldots & \psi_R(\omega_1, \omega_2)_R \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(R) \end{bmatrix}
$$

$$(3.11)$$

26

In vector-matrix notation, Eq. (3.11) becomes

$$\mathbf{H} = \mathbf{Ba}, \tag{3.12}$$

in which $\mathbf{B}$ is of dimension $R \times R$. The sequence $a(n_1, n_2)$ results from solving the relation

$$\mathbf{a} = \mathbf{B}^{-1}\mathbf{H}. \tag{3.13}$$

The final step in the process is to compute the filter impulse response $h(n_1, n_2)$ using the relationship of Eq. (3.6). The resulting impulse response has an $N_1 \times N_2$ region of support. Equation (3.4) suggests a transfer function of the form

$$
\begin{aligned}
H(z_1, z_2) = & h(0,0) + \sum_{n_1=1}^{M_1} h(n_1,0)(z_1^{n_1} + z_1^{-n_1}) + \sum_{n_2=1}^{M_2} h(0,n_2)(z_2^{n_2} + z_2^{-n_2}) \\
& + \sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} h(n_1,n_2)(z_1^{n_1} + z_1^{-n_1})(z_2^{n_2} + z_2^{-n_2}),
\end{aligned}
\tag{3.14}
$$

which can be used in implementing a 2-D filter structure with a minimal number of gains.

This same approach can also be employed if other than rectangular impulse response regions of support are specified. Again, a system of linear equations of the form of Eq. (3.5) is solved, but the limits on the summations are altered to reflect the desired region of support for $h(n_1, n_2)$.

To summarize the arbitrary-sampling design process, the following steps are performed:

(1) Fix the impulse response region of support, and hence, the filter order.

(2) Specify the desired frequency response at selected sample points on the region $K = \{(\omega_1, \omega_2) : 0 \le \omega_1 \le \pi; 0 \le \omega_2 \le \pi\}$. The number of specified samples must equal the number of independent impulse response values (i.e., those which are not fixed by the symmetry conditions). The

27

locations of samples within the region is unconstrained except for the condition that each sample location be distinct.

(3) Solve a system of linear equations to obtain the resulting impulse response $h(n_1, n_2)$.

B. EXISTENCE AND UNIQUENESS

There is a fundamental deficiency in the arbitrary-sampling approach to the 2-D FIR frequency-sampling filter design problem. For specified frequency response samples placed arbitrarily in region K, there is no guarantee that a unique filter possessing a number of independent impulse response samples equal to the number of interpolating conditions, and which satisfies each of those interpolating conditions, exists. This is a problem inherent to interpolation in two dimensions which does not arise in the one-dimensional case.

For a unique solution to the system of equations indicated by Eq. (3.12) to exist, the basis functions $\{\psi_i(\omega_1, \omega_2)\}$ must satisfy the Haar condition. A set of vectors in n-space satisfies the Haar condition if every set of n of them is linearly independent [Ref. 7: p. 45]. Therefore, to satisfy the Haar condition, any set of R characteristic vectors defined on R distinct points in region K of the (ω_1, ω_2) plane must be linearly independent. For the arbitrary-sampling method developed in this chapter, a characteristic vector $\{\Psi((\omega_1, \omega_2)_i)\}$ associated with a particular sample location $(\omega_1, \omega_2)_i$ can be defined as

$$\Psi((\omega_1, \omega_2)_i) = [\psi_1((\omega_1, \omega_2)_i) \quad \psi_2((\omega_1, \omega_2)_i) \quad \ldots \quad \psi_R((\omega_1, \omega_2)_i)]^T . \quad (3.15)$$

It is known that for a region K' of the two-dimensional frequency plane, there are no nontrivial sets of two-dimensional basis functions which satisfy the Haar condition [Ref. 8: p. 494]. Therefore, there is no guarantee that the sets of characteristic vectors formed by evaluating Eq. (3.15) at arbitrarily selected sample points in

28

region K of the (ω_1, ω_2) plane are linearly independent. Since the square matrix **B** in Eq. (3.12) can be expressed in terms of the characteristic vectors as

$$\mathbf{B} = [\; \boldsymbol{\Psi}((\omega_1, \omega_2)_1) \quad \boldsymbol{\Psi}((\omega_1, \omega_2)_2) \quad \ldots \quad \boldsymbol{\Psi}((\omega_1, \omega_2)_R)\;]^T, \qquad (3.16)$$

and since the characteristic vectors may be linearly dependent, **B** may be singular and a solution to Eq. (3.13) may not exist. Such a case is termed a degeneracy, and illustrates a basic problem associated with the arbitrary-sampling method. As can be inferred from the form of the basis functions, the occurrence of a degeneracy is dependent solely upon the location of samples; the specified frequency response values at the sample points do not affect this.

A simple example illustrates how a degeneracy can arise. The frequency response of a filter is specified at four sample points in the (ω_1, ω_2) plane:

1. $H(0, 0.6\pi) = 1$ 2. $H(0.4\pi, 0) = 1$

3. $H(0.4\pi, \pi) = 0$ 4. $H(\pi, 0.6\pi) = 0$

These sample locations are shown in Figure 3.2.

The parameters M_1 and M_2 are both set equal to 1 so that the number of samples in the sequence $a(n_1, n_2)$ equals the number of frequency samples. Substitution into Eq. (3.11) yields

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & \cos 0.6\pi & 1 & \cos 0.6\pi \\ 1 & 1 & \cos 0.4\pi & \cos 0.4\pi \\ 1 & -1 & \cos 0.4\pi & -\cos 0.4\pi \\ 1 & \cos 0.6\pi & -1 & -\cos 0.6\pi \end{bmatrix} \begin{bmatrix} a(0,0) \\ a(0,1) \\ a(1,0) \\ a(1,1) \end{bmatrix}. \qquad (3.17)$$

Some calculations reveal that the square matrix above has a determinant of zero; therefore, it is not invertible and no unique solution for $a(n_1, n_2)$ exists.

To determine whether degenerate cases present a serious shortcoming of this method, it is of interest to determine how often degeneracies are likely to arise. To investigate this, an APL computer program was developed to count the number of
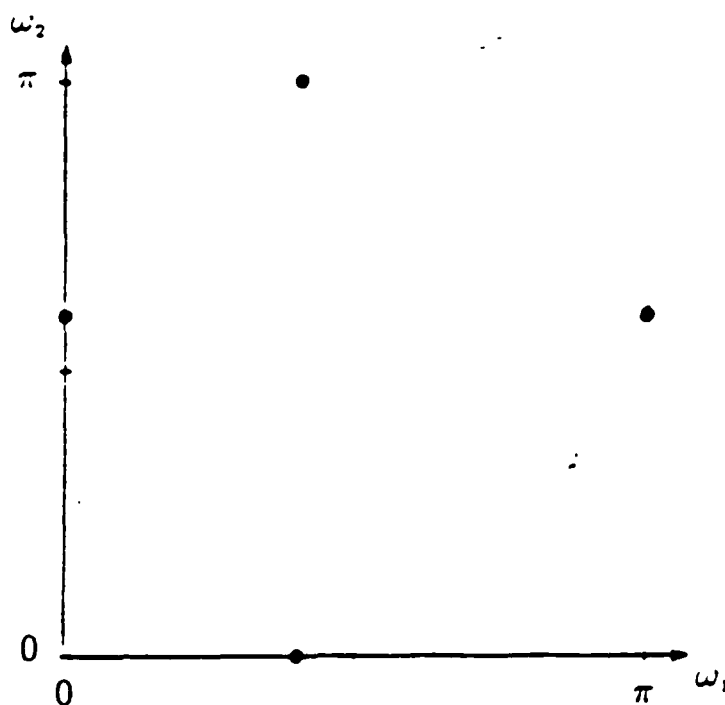
Figure 3.2. Arrangement of Samples for Design Example.

degenerate cases arising from the arbitrary selection of sample points $\{(\omega_1, \omega_2)_i\}$
on region K of the (ω_1, ω_2) plane. A total of R samples were arbitrarily placed
at distinctly different vertices of a grid of specified density on region K for each
case. The placement of samples was performed in a pseudo-random fashion using
the APL "roll" function. The density of the grid, in effect, is determined by the
number of decimal places carried in the frequency values. In the analysis, two
parameters were varied: the dimensions of the specified filter impulse response,
and the grid density. For each situation, one hundred separate sets of arbitrary
samples were investigated. The program then determined the number of cases in
which the sampling geometry resulted in a degeneracy. Due to the large number of
computations involved, analysis was limited to filters with square impulse response
regions of support only. The results are listed in Table 3-1.

30

These results support two concepts. First, for a fixed grid density, as the order of the specified filter increases (and hence the number of specified frequency samples), the likelihood of the occurrence of a degeneracy increases. Secondly, for a fixed filter order, as the density of the grid on which samples are placed increases, the likelihood of a degeneracy occurring decreases. For filters of order 17×17 or less, the probability of a degenerate case is minimal provided frequencies are specified to at least the nearest 0.05 radians.

TABLE 3-1

NO. OF DEGENERACIES OCCURRING IN 100 TEST CASES

FILTER ORDER	R	GRID DENSITY (RADIANS)		
		0.2	0.1	0.05
5×5	9	5	1	0
9×9	25	10	0	0
13×13	49	17	1	0
17×17	81	56	2	0

It is significant to note that for the above tests, frequency samples were chosen in an arbitrary manner. However, when designing a filter to approximate a certain desirable frequency response characteristic, the selection of points is likely to be done in an unconstrained, but not arbitrary, manner. Certain patterns in sample selection, such as selecting octant symmetric sample points where $(\omega_1, \omega_2)_i = (\omega_2, \omega_1)_j$, have a greater tendency to introduce linear dependencies within the set of characteristic vectors than when sample points are selected at random. This concept is difficult to show in a controlled experiment; however, in

31

the course of this research, degeneracies arose when attempting filter designs much more often than the arbitrarily placed sample test results would indicate.

Such a degeneracy did arise in the following design example. In this case, a lowpass filter is desired with a circularly symmetric frequency response such that it possesses unity gain in the passband region $\{0 \leq \sqrt{\omega_1^2 + \omega_2^2} \leq 0.2\pi\}$ and zero gain in the stopband region $\{0.4\pi \leq \sqrt{\omega_1^2 + \omega_2^2} \leq \pi\}$. The region of support for the impulse response is specified to be 17×17, implying $M_1 = M_2 = 8$. Therefore, a total of $R = (M_1 + 1)(M_2 + 1) = 81$ frequency samples are to be specified. Figure 3.3 shows the location of 81 points at which the desired frequency response is to be sampled initially.
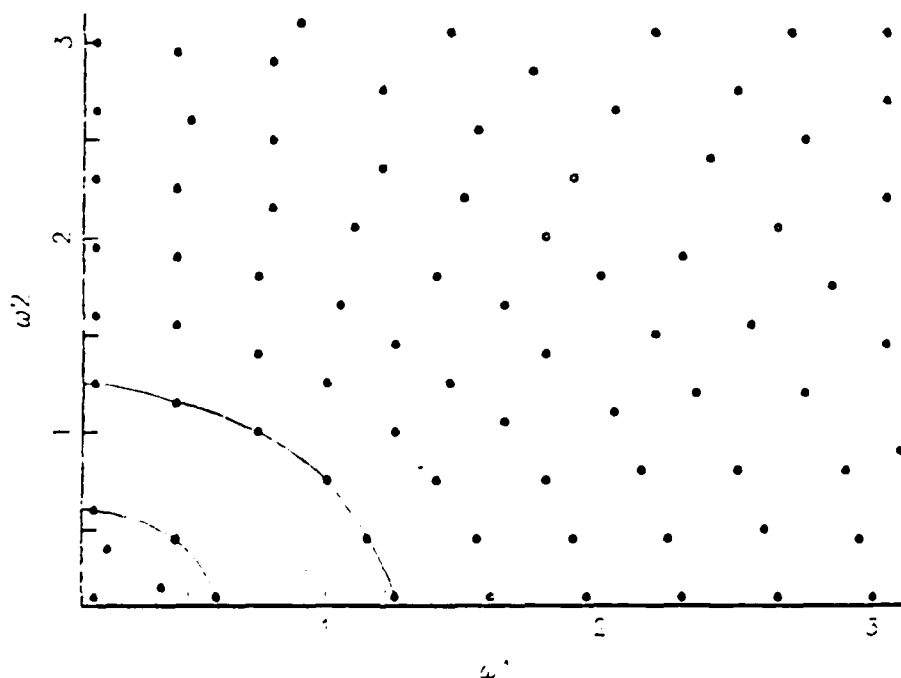
Figure 3.3. Arrangement of Samples for Lowpass Filter
Design, Case 1.

32

Note that the arbitrary-sampling method does not require any samples to be placed in the transition (don't care) region. This allows the concentration of more samples in the passband and stopband regions. This particular choice of sample locations, however, was found to result in a degeneracy when a computer solution was attempted.

Following this degenerate result, the sampling geometry was altered by slight displacement of the locations of six of the frequency samples. The new sampling geometry is illustrated in Figure 3.4, with the circled points indicating the sample locations which differ from the degenerate case. The resulting filter frequency response is shown in Figure 3.5.
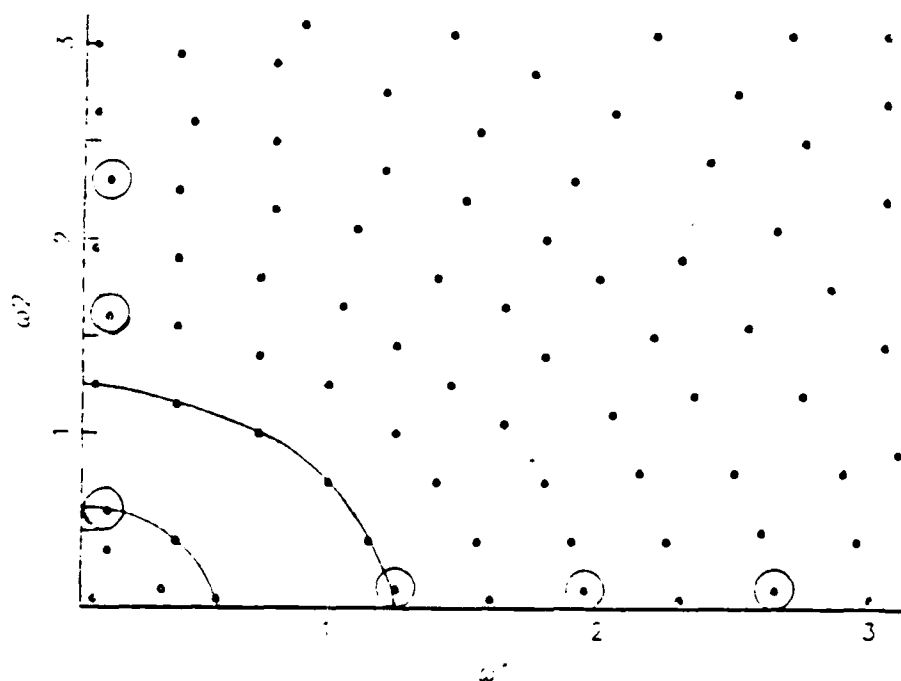
Figure 3.4. Arrangement of Samples for Lowpass Filter
Design, Case 2.

33
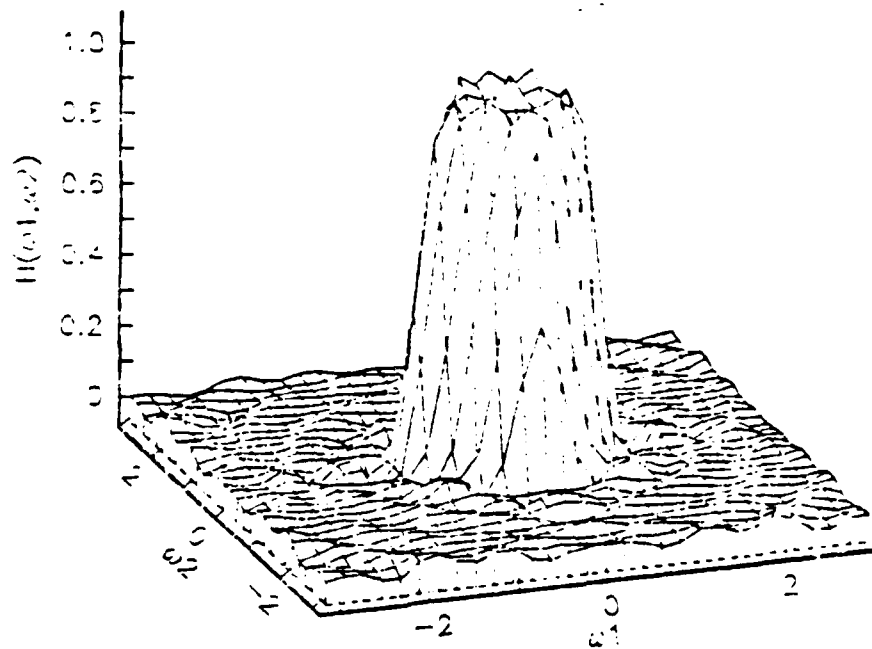
Figure 3.5. Frequency Response of Lowpass Filter
Design, Case 2.

Although the result may seem to be perfectly acceptable, the system of equations involved in the solution can be considered to be ill-conditioned. Slight perturbation of the position of a single sample point, as shown by the circled point in Figure 3.6, results in the unacceptable, excessively rippled response indicated by Figure 3.7.

The above example demonstrates a negative aspect of the arbitrary-sampling approach. Although degeneracies, as defined, may seldom occur, ill-conditioned cases in which the determinant of matrix B in Eq. (3.12) is very small , but nonzero, may arise much more frequently. While specifying the frequencies ω_1 and ω_2 at each sample point with a high degree of precision (analogous to placing samples on a very dense grid) may assure a very minimal probability of a degeneracy, it does not guarantee that the resulting system will not be ill-conditioned.
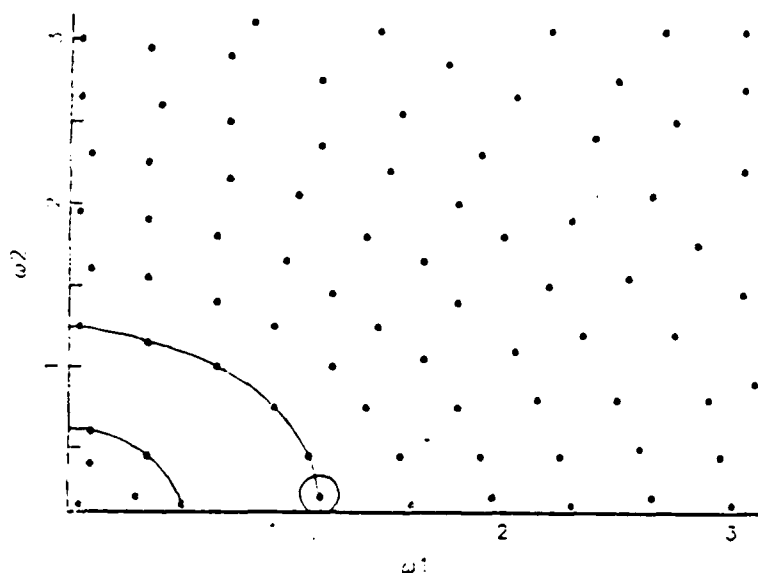
34

Figure 3.6. Arrangement of Samples for Lowpass Filter
Design, Case 3.

C. COMPUTATIONS

A further negative aspect of this arbitrary-sampling method is the number
of computations involved in determining filter parameters. While the uniform-
sampling method of Chapter II was shown to be reasonably efficient in computation,
the arbitrary-sampling method involves the solution of a rather large system of
linear equations. This solution requires the inversion of an $R \times R$ matrix. This
inversion dominates the computation process. It is well known that inverting an
$N \times N$ matrix requires on the order of N^3 floating point operations, where a floating
point operation consists of one multiplication and one addition. Since, for a filter
with an $N_1 \times N_2$ impulse response region of support,

$$
\begin{aligned}
R = (M_1 + 1)(M_2 + 1) &= \left(\frac{N_1 - 1}{2} + 1\right)\left(\frac{N_2 - 1}{2} + 1\right) \\
&= \frac{1}{4}(N_1 + 1)(N_2 + 1)
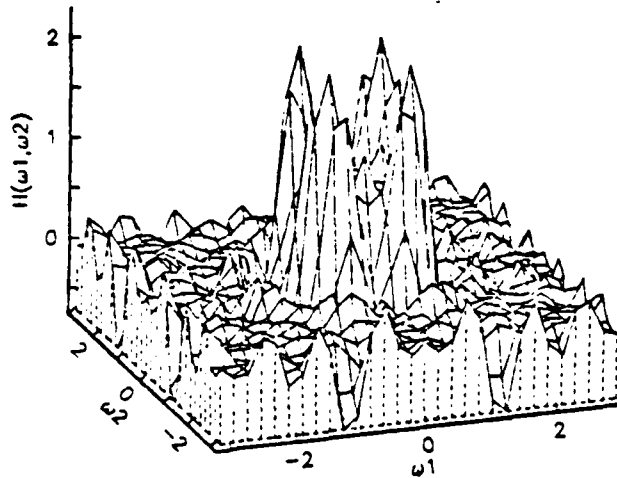\end{aligned} \tag{3.18}
$$

35

Figure 3.7. Frequency Response of Lowpass Filter
Design, Case 3.

the order of the number of floating point operations involved in the arbitrary-sampling design process can be represented as

$$C_{arb} \simeq \frac{1}{64} N_1^3 N_2^3 . \tag{3.19}$$

For the case in which $N_1 = N_2 = N$, this reduces to the following:

$$C_{arb} \simeq \frac{1}{64} N^6 . \tag{3.20}$$

Thus, the arbitrary sampling method represents an approach which, while quite flexible, is computationally intensive and fails to provide guarantee of a unique solution. In the next chapter, an alternative nonuniform sampling approach is presented which, by placing some mild constraints on frequency sample location, always provides a unique design solution and is much more computationally efficient.

IV. A NOVEL 2-D NONUNIFORM FREQUENCY-SAMPLING APPROACH

Chapters II and III presented two fundamentally different approaches to 2-D FIR frequency-sampling filter design. The uniform-sampling method possessed no flexibility in the location of frequency samples, yet was computationally efficient in determining filter parameters. The arbitrary sampling method, however, had few constraints on sample location but was computationally intensive. The approach described in this chapter represents a compromise. It is more constrained than the arbitrary sampling method, but has much more freedom in sample placement than for the uniform-sampling method. While not as efficient as the IDFT approach, this method requires significantly fewer computations when compared with the arbitrary sampling method. Additionally, existence and uniqueness of a solution for the impulse response is guaranteed, provided simple constraints on sample locations are met.

A. APPROACH

For convenience, the design method introduced in this chapter will be referred to as the "nonuniform sampling" method, although the arbitrary sampling method uses nonuniformly spaced samples as well. This method borrows ideas from the other two approaches. As in the arbitrary-sampling case, quadrant symmetry is specified. As was shown in Chapter III, the frequency response of a 2-D zero-phase FIR filter with an $N_1 \times N_2$ impulse response region of support can be represented as

$$H(\omega_1, \omega_2) = \sum_{n_1=0}^{M_1} \sum_{n_2=0}^{M_2} a(n_1, n_2) \cos \omega_1 n_1 \cos \omega_2 n_2, \tag{4.1}$$

37

where

$$a(n_1, n_2) = \begin{cases} h(n_1, n_2) & n_1 = 0, n_2 = 0 \\ 2h(n_1, n_2) & n_1 \neq 0, n_2 = 0 \\ 2h(n_1, n_2) & n_1 = 0, n_2 \neq 0 \\ 4h(n_1, n_2) & n_1 \neq 0, n_2 \neq 0 \end{cases}, \qquad (4.2)$$

with $M_1 = (N_1 - 1)/2$ and $M_2 = (N_2 - 1)/2$. In the arbitrary sampling method, the sequence $a(n_1, n_2)$ was represented as a column vector, the basis functions $\psi_i(\omega_1, \omega_2)$ were specified, and a system of equations of the form of Eq. (4.1) was solved directly. In the nonuniform sampling method, constraints placed on the sampling geometry result in Eq. (4.1) becoming a separable system. The system is then solved by a two-step process analogous to the row-column decomposition used in the uniform sampling method of Chapter II. The resulting systems of linear equations contain one-dimensional basis functions only; hence, linear dependencies associated with two-dimensional basis functions are not of concern. This results in assurance of a unique solution to the filter design problem.

As in the arbitrary sampling method, samples are to be placed in region $K = \{(\omega_1, \omega_2) : 0 \leq \omega_1 \leq \pi; 0 \leq \omega_2 \leq \pi\}$ of the two-dimensional frequency plane. As mentioned above, certain constraints on the location of frequency samples are required. Suppose any $(M_1 + 1)$ distinct values of ω_1 are chosen on the interval $0 \leq \omega_1 \leq \pi$. These values are denoted as $\{\omega_{1k} : k = 0, 1, \ldots, M_1\}$. Then for each ω_{1k}, suppose any $(M_2 + 1)$ distinct values of ω_2 are chosen such that $0 \leq \omega_2 \leq \pi$. These values are denoted as $\{\omega_{2l}(k) : l = 0, 1, \ldots, M_2\}$. Then a unique solution for $a(n_1, n_2)$ can be obtained from

$$H(\omega_{1k}, \omega_{2l}(k)) = \sum_{n_1=0}^{M_1} \sum_{n_2=0}^{M_2} a(n_1, n_2) \cos \omega_{1k} n_1 \cos \omega_{2l}(k) n_2;$$
$$0 \leq k \leq M_1, \quad 0 \leq l \leq M_2. \qquad (4.3)$$

Such a sampling geometry is shown in Figure 4.1 for the case $M_1 = M_2 = 2$. Notice that the selection of sample points is quite flexible. The values ω_{10}, ω_{11},

38

and ω_{12} can be any arbitrary, distinct samples of ω_1 in the region. Additionally, $\omega_{20}(k)$, $\omega_{21}(k)$, and $\omega_{22}(k)$ represent any arbitrary, distinct samples of ω_2 for each k, $k = 0, 1, 2$. This is far less constrained than placing the samples on a uniform Cartesian grid.
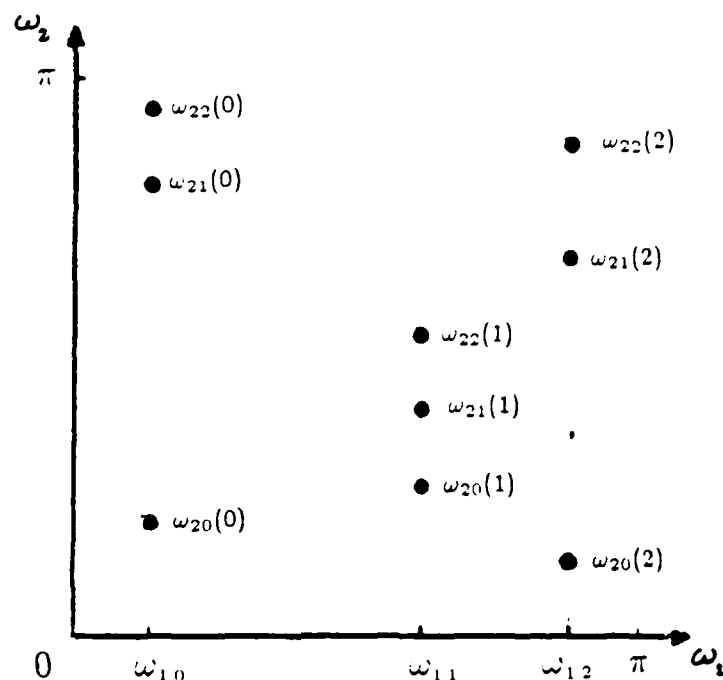

Figure 4.1. Nonuniform Sampling Example.

Given the arrangement of samples described above, the sequence $a(n_1, n_2)$ can be found as follows. First, Eq. (4.1) is arranged as

$$H(\omega_1, \omega_2) = \sum_{n_2=0}^{M_2} \Big[\sum_{n_1=0}^{M_1} a(n_1, n_2) \cos \omega_1 n_1 \Big] \cos \omega_2 n_2. \qquad (4.4)$$

A new function $g(\omega_1, n_2)$ is defined as

$$g(\omega_1, n_2) = \sum_{n_1=0}^{M_1} a(n_1, n_2) \cos \omega_1 n_1. \qquad (4.5)$$

39

The frequency response can then be expressed as

$$H(\omega_1, \omega_2) = \sum_{n_2=0}^{M_2} g(\omega_1, n_2) \cos \omega_2 n_2. \tag{4.6}$$

For a particular value of ω_1, namely ω_{1k}, the response can be written as

$$H(\omega_{1k}, \omega_2) = \sum_{n_2=0}^{M_2} g(\omega_{1k}, n_2) \cos \omega_2 n_2. \tag{4.7}$$

Now, for each $\omega_{2l}(k)$, $l = 0, 1, \ldots, M_2$, the above expression becomes

$$H(\omega_{1k}, \omega_{2l}(k)) = \sum_{n_2=0}^{M_2} g(\omega_{1k}, n_2) \cos \omega_{2l}(k) n_2. \tag{4.8}$$

Thus, the first step in the row-column decomposition is to solve a system of $(M_2 + 1)$ linear equations of the form of Eq. (4.8). In matrix form, this becomes

$$\begin{bmatrix} H(\omega_{1k}, \omega_{20}(k)) \\ H(\omega_{1k}, \omega_{21}(k)) \\ \vdots \\ H(\omega_{1k}, \omega_{2M_2}(k)) \end{bmatrix} = \begin{bmatrix} 1 & \cos \omega_{20}(k) & \ldots & \cos M_2 \omega_{20}(k) \\ 1 & \cos \omega_{21}(k) & \ldots & \cos M_2 \omega_{21}(k) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos \omega_{2M_2}(k) & \ldots & \cos M_2 \omega_{2M_2}(k) \end{bmatrix} \begin{bmatrix} g(\omega_{1k}, 0) \\ g(\omega_{1k}, 1) \\ \vdots \\ g(\omega_{1k}, M_2) \end{bmatrix}. \tag{4.9}$$

In vector-matrix notation, Eq. (4.9) is expressed as

$$\mathbf{H}_k = \mathbf{A} \mathbf{g}_k, \tag{4.10}$$

where $\mathbf{H}_k$ and $\mathbf{g}_k$ are $(M_2 + 1) \times 1$ column vectors, and $\mathbf{A}$ is an $(M_2 + 1) \times (M_2 + 1)$ square matrix. The vector $\mathbf{g}_k$ is then computed using the relation

$$\mathbf{g}_k = \mathbf{A}^{-1} \mathbf{H}_k. \tag{4.11}$$

This entire process is performed for each k, $k = 0, 1, \ldots, M_1$. Therefore, this step involves a total of $(M_1 + 1)$ systems of $(M_2 + 1)$ linear equations each.

Once the column vectors $\mathbf{g}_k$ have been computed for all values of k, a matrix $\mathbf{G}$ of size $(M_2 + 1) \times (M_1 + 1)$ can be formed such that

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 & \mathbf{g}_1 & \cdots & \mathbf{g}_{M_1} \end{bmatrix}. \tag{4.12}$$

40

Now the second step of the row-column decomposition can be performed. From Eq. (4.5), for a particular ω_{1k},

$$g(\omega_{1k}, n_2) = \sum_{n_1=0}^{M_1} a(n_1, n_2) \cos \omega_{1k} n_1. \tag{4.13}$$

For a particular n_2, a system of linear equations can be formed by evaluating Eq. (4.13) for each k, $k = 0, 1, \ldots, M_1$. In matrix form, this becomes

$$
\begin{bmatrix}
g(\omega_{10}, n_2) \\
g(\omega_{11}, n_2) \\
\vdots \\
g(\omega_{1M_1}, n_2)
\end{bmatrix}
=
\begin{bmatrix}
1 & \cos \omega_{10} & \ldots & \cos M_1 \omega_{10} \\
1 & \cos \omega_{11} & \ldots & \cos M_1 \omega_{11} \\
\vdots & \vdots & \ddots & \vdots \\
1 & \cos \omega_{1M_1} & \ldots & \cos M_1 \omega_{1M_1}
\end{bmatrix}
\begin{bmatrix}
a(0, n_2) \\
a(1, n_2) \\
\vdots \\
a(M_1, n_2)
\end{bmatrix}. \tag{4.14}
$$

Equation (4.14) can be expressed in vector-matrix notation as

$$\mathbf{\Gamma}_{n_2} = \mathbf{B}\mathbf{a}, \tag{4.15}$$

where $\mathbf{\Gamma}_{n_2}$ is a $(M_1 + 1) \times 1$ column vector formed from the elements of the n_2th row of $\mathbf{G}$, $\mathbf{B}$ is a $(M_1 + 1) \times (M_1 + 1)$ square matrix and $\mathbf{a}$ is a $(M_1 + 1) \times 1$ column vector. Thus $\mathbf{a}$ is found from the relation

$$\mathbf{a} = \mathbf{B}^{-1}\mathbf{\Gamma}_{n_2}. \tag{4.16}$$

The vector $\mathbf{a}$ represents the values of $a(n_1, n_2)$ for a specified value of n_2. To obtain the entire $(M_1 + 1) \times (M_2 + 1)$ sequence $a(n_1, n_2)$, the process described by Eq. (4.16) is repeated for each value of n_2, $n_2 = 0, 1, \ldots, M_2$. Therefore, a total of $(M_2 + 1)$ systems containing $(M_1 + 1)$ linear equations each are involved in the second step of the method. Finally, the impulse response of the resulting filter can be evaluated using the relationship of Eq. (4.2).

The following design example illustrates how this method can be employed. For this example, a 2-D circularly symmetric lowpass filter characteristic is desired

41

with unity gain in the passband region $\{(\omega_1, \omega_2) : 0 \leq \sqrt{\omega_1^2 + \omega_2^2} \leq 0.2\pi\}$ and zero gain in the stopband region $\{(\omega_1, \omega_2) : 0.4\pi \leq \sqrt{\omega_1^2 + \omega_2^2} \leq \pi\}$. The region of support for the impulse response is specified to be 17×17. The arrangement of frequency samples shown in Figure 4.2 is to be used. Note that the sampling is not uniform. When samples are placed in the transition region, a linear slope characteristic between the two bands is used in specifying the desired frequency response value at those samples, as was done in the uniform sampling approach. Contour and perspective plots of the frequency response for the resulting filter are shown in Figure 4.3.
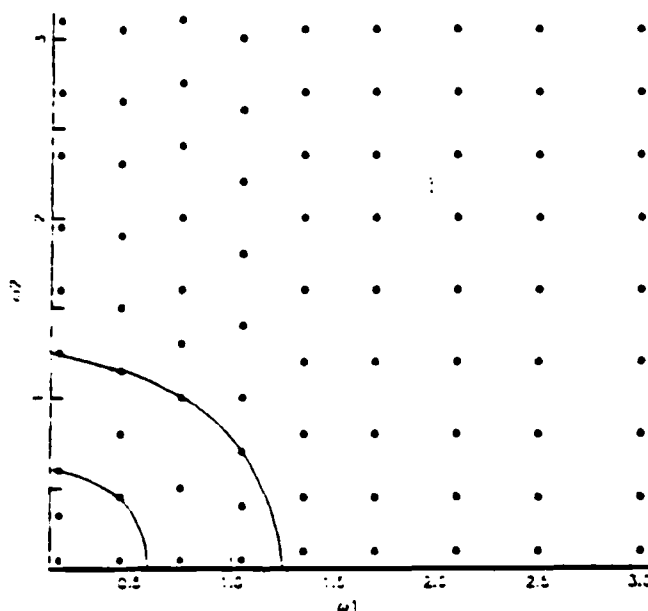
Figure 4.2. Arrangement of Samples for 17×17 Lowpass Filter.

One rather obvious extension of the nonuniform sampling method involves again specifying an $N_1 \times N_2$ impulse response region of support, but in this case performing the interpolation in the first step of the process on ω_1. Suppose any $(M_2 + 1)$ distinct values of ω_2 are selected on the interval $0 \leq \omega_2 \leq \pi$. These are denoted as $\{\omega_{2l}, \ l = 0, 1, \ldots, M_2\}$. For each value of ω_{2l}, any $(M_1 + 1)$ distinct

42
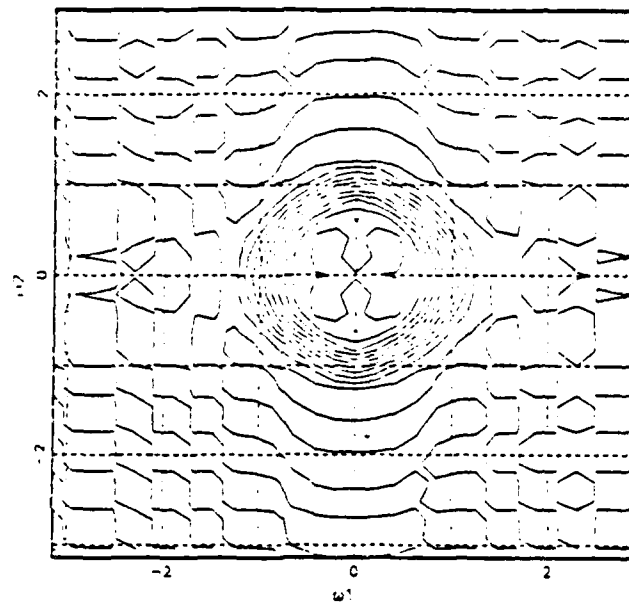
(a)
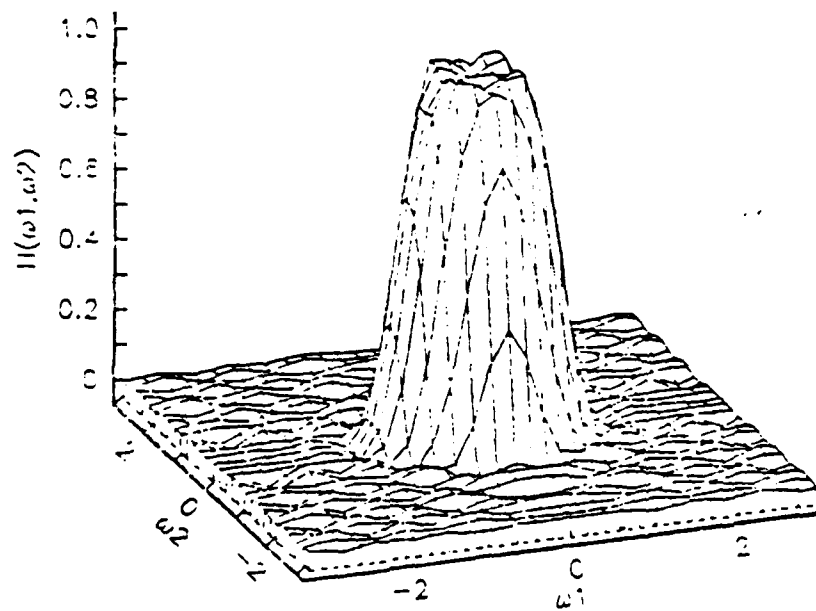
(b)

Figure 4.3. Frequency Response of 17 × 17 Lowpass Filter.

43

values of ω_1 are chosen on the interval $0 \leq \omega_1 \leq \pi$. These values are denoted as $\{\omega_{1k}(l),\ k = 0,1,\ldots,M_1\}$. Such an arrangement of selected samples is illustrated in Figure 4.4 for the case $M_1 = M_2 = 2$. Now suppose $H(\omega_1,\omega_2)$ is specified at each selected sample $(\omega_{1k}(l),\omega_{2l})$ for $k = 0,1,\ldots,M_1$ and $l = 0,1,\ldots,M_2$.
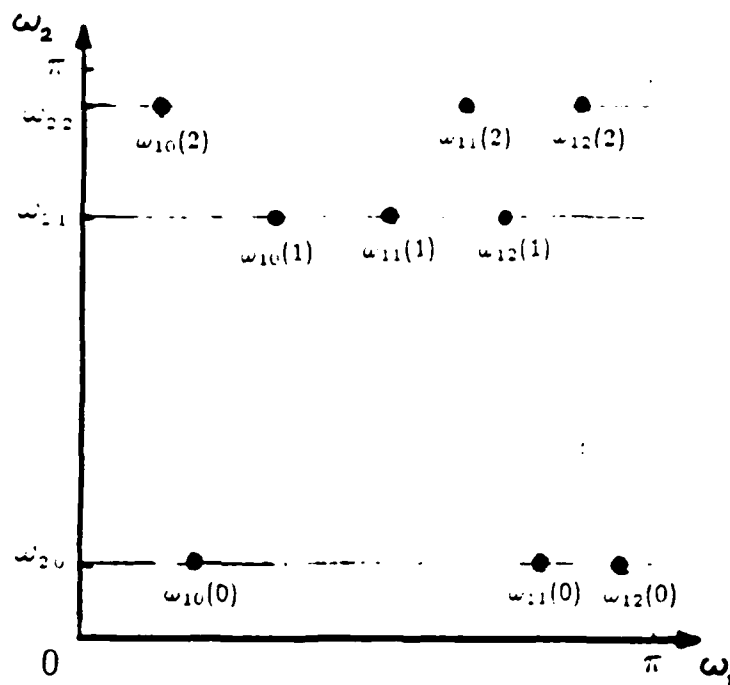

Figure 4.4. Alternate Nonuniform Sampling Approach.

Then there is a unique solution for $a(n_1,n_2)$ which can be obtained from

$$H(\omega_{1k}(l),\omega_{2l}) = \sum_{n_1=0}^{M_1} \sum_{n_2=0}^{M_2} a(n_1,n_2)\cos\omega_{1k}(l)n_1 \cos\omega_{2l}n_2;$$
$$0 \leq k \leq M_1, \quad 0 \leq l \leq M_2.$$

(4.17)

The approach to the solution of the above problem and proof of its existence and uniqueness can be performed in a straightforward manner analogous to that of the basic approach.

44

It is worthy to note that the sampling geometry required by the uniform sampling method of Chapter II is, in reality, a special case of the general nonuniform sampling form discussed here. Therefore, the results of the uniform sampling method can always be achieved by constraining the location of samples to the uniform Cartesian grid. However, the method of solution is less efficient than the IDFT approach and so would probably not be used for this particular case.

B. EXISTENCE AND UNIQUENESS

A significant aspect of the nonuniform sampling method is that a unique solution matching the specified frequency response at each sample point is guaranteed, provided samples are located as described. Proof of existence and uniqueness of a sequence $h(n_1, n_2)$ which satisfies the interpolating conditions follows.

The first step in the nonuniform sampling approach involves solution of $(M_1 + 1)$ systems of linear equations of the form of Eq. (4.9) or Eq. (4.10). A unique solution exists for each of these systems. Recall that Eq. (4.9) resulted from applying the interpolating conditions to Eq. (4.6). Through use of the Chebyshev polynomials, Eq. (4.6) can be expressed as

$$H(\omega_1, \omega_2) = \sum_{n_2=0}^{M_2} c(\omega_1, n_2)(\cos \omega_2)^{n_2}, \qquad (4.18)$$

where a unique relation exists between the coefficients $g(\omega_1, n_2)$ and $c(\omega_1, n_2)$. Through the subtitution $x = \cos \omega_2$, the above becomes

$$H(\omega_1, x) = \sum_{n_2=0}^{M_2} c(\omega_1, n_2)x^{n_2}. \qquad (4.19)$$

The set $\{1, x, x^2, \ldots, x^{M_2}\}$ is known to be unisolvent on any interval [a,b] [Ref. 9: p. 31]. Therefore, a unique interpolation polynomial of the form of Eq. (4.19) can be found given any $(M_2 + 1)$ distinct values of x. From the relation $x_l = \cos \omega_{2l}$,

45

as long as distinct $\omega_{2l}(k)$ samples are specified on the interval $\{0 \leq \omega_2 \leq \pi\}$, the corresponding x_l will be distinct. Due to the unique relationship between Eqs. (4.6), (4.18), and (4.19), the interpolation implied by Eq. (4.10) will yield a unique solution for $\mathbf{g}_k$, $k = 0, 1, \ldots, M_1$. This implies that the matrix $\mathbf{G}$ defined in Eq. (4.12) will be unique.

The second step of the nonuniform sampling approach involves solution of $(M_2 + 1)$ systems of linear equations of the form of Eq. (4.15). Each row of matrix $\mathbf{B}$ in Eq. (4.15) forms a vector of Chebyshev polynomials which form a unisolvent set. Therefore, $\mathbf{B}$ is invertible. Since a unique matrix $\mathbf{G}$ was found in the first step, the column vectors $\boldsymbol{\Gamma}_{n_2}$ are unique for each $n_2 = 0, 1, \ldots, M_2$. Therefore, a unique solution for $a(n_1, n_2)$ exists which satisfies the overall interpolation problem. From the relationship indicated by Eq. (4.2), this implies existence and uniqueness for the resulting impulse response $h(n_1, n_2)$.

Ill-conditioned results normally occur in interpolation problems involving one variable when one sample is located in close proximity to another. Since the nonuniform sampling method uses a two-step process involving one-dimensional interpolating functions, the same result can be expected to hold here. Provided samples are not placed too close together, the resulting filter design should not be ill-conditioned.

C. COMPUTATIONS

While the nonuniform sampling method retains less flexibility than the arbitrary sampling method of Chapter III, it is more computationally efficient for filters of the same order. As in the analysis for the arbitrary sampling case, inverting an $N \times N$ matrix is assumed to require on the order of N^3 floating point operations and additions, and inversion of the matrix is assumed to dominate the

46

computations required in solving a system of linear equations. A brief analysis of the number of operations required for the nonuniform sampling method follows.

For a filter with a specified impulse response region of support of size $N_1 \times N_2$, step one of the design process requires solution of a total of $(M_1 + 1)$ systems of $(M_2 + 1)$ linear equations each. The number of floating point operations involved in this step is therefore on the order of $(M_1 + 1)(M_2 + 1)^3$.

Step two of the design process requires solution of a total of $(M_2 + 1)$ systems of $(M_1 + 1)$ linear equations each. The number of required floating point operations in this step, then, is on the order of $(M_2 + 1)(M_1 + 1)^3$.

The total number of floating point operations for the entire nonuniform frequency-sampling 2–D FIR filter design method is the combined amount from the above two steps. The total number, therefore, is on the order of

$$
\begin{aligned}
C_{\mathrm{nu}} &\simeq (M_1 + 1)(M_2 + 1)^3 + (M_1 + 1)^3(M_2 + 1) \\
&= \left(\frac{N_1 - 1}{2} + 1\right)\left(\frac{N_2 - 1}{2} + 1\right)^3 + \left(\frac{N_1 - 1}{2} + 1\right)^3\left(\frac{N_2 - 1}{2} + 1\right) \\
&= \frac{1}{16}(N_1 + 1)(N_2 + 1)^3 + \frac{1}{16}(N_1 + 1)^3(N_2 + 1) \\
&\simeq \frac{1}{16}(N_1 N_2^3 + N_1^3 N_2).
\end{aligned}
\tag{4.20}
$$

For the particular case in which $N_1 = N_2 = N$, Eq. (4.20) reduces to

$$
C_{\mathrm{nu}} \simeq \frac{1}{8}N^4.
\tag{4.21}
$$

Thus, the nonuniform sampling method is much more computationally efficient than the arbitrary sampling method, yet is still quite flexible while providing guarantee of a unique design solution.

V. A COMPARISON OF FREQUENCY-SAMPLING DESIGN METHODS THROUGH DESIGN EXAMPLES

Up to this point the theory behind three fundamentally different approaches to the frequency-sampling design of 2–D FIR filters has been developed. With knowledge of these approaches, it is of interest to investigate the results of applying each of the methods to some practical design problems. In this chapter, some design problems are formulated and the results from each of the methods are compared in terms of performance, computations, and flexibility.

A. APPROACH

In order to compare the three basic design methods discussed in this thesis, three different design problems are described. To achieve a valid comparison, several parameters were fixed for each design. The region of support for the filter impulse response in all cases was specified to be 17×17. A square region of support is most commonly used in 2–D filter design and avoids some of the problems discussed in the latter part of Chapter IV. The specified dimensionality of the impulse response is large enough to ensure reasonable results but not so large as to make implementation impractical.

Each design method requires specification of a desired frequency response $H_d(\omega_1, \omega_2)$. For each design method, $H_d(\omega_1, \omega_2)$ is fixed, with unity gain throughout the passband, zero gain throughout the stopband, and a linear slope characteristic in the transition band(s). The frequency response value assigned to each sample point in the two-dimensional frequency domain is the value of $H_d(\omega_1, \omega_2)$ at that point.

48

For each design problem, circularly symmetric filters are prescribed. Although more general symmetries are allowable for each method, circularly symmetric designs adequately illustrate the salient characteristics of the methods.

A few other points are of note. For these design examples, optimimization techniques were not utilized in locating frequency samples for any of the methods. Frequency sample locations were selected strictly on a trial–and–error basis by the human designer for the arbitrary and nonuniform sampling methods. The resulting frequency responses of the filters are compared qualitatively in terms of passband and stopband ripple, transition band width, and circularity. No quantitative comparisons are made in terms of best equi-ripple approximation to an ideal response characteristic, etc. It is well known that a better quality filter can usually be obtained by increasing the number of terms in the impulse response (i.e. increasing the filter order). In comparing frequency-sampling methods with the same support for the impulse response in each case, substantially better filters are not expected to result from one particular method; rather, it is of interest to investigate how design method flexibility can be used to enhance certain desirable characteristics in the resulting filter frequency response.

B.   DESIGN EXAMPLES

Three different design problems are now presented. The three frequency-sampling design methods are applied to each and a comparative analysis of the resulting filter frequency responses is performed.

1.   Lowpass Filter No. 1

For this example, a circularly symmetric 2–D lowpass filter frequency response is desired such that it has unity gain in the passband region $\{0 \leq \sqrt{\omega_1^2 + \omega_2^2} \leq 0.4\pi\}$ and zero gain in the stopband region $\{0.6\pi \leq \sqrt{\omega_1^2 + \omega_2^2} \leq \pi\}$. First, a 17×17

49

filter was designed to match the above characteristics using the uniform sampling method. Contour and perspective plots of the filter frequency response are shown in Figure 5.1. Next, a 17×17 filter was designed using the nonuniform sampling method with the arrangement of samples shown in Figure 5.2. The resulting filter frequency response is illustrated in Figure 5.3. Finally, the arbitrary sampling method was applied to this design problem. The selection of samples was as shown in Figure 5.4, and the frequency response of the resulting filter can be seen in Figure 5.5.

In comparing the results of the three methods, the quality of the resulting filters varies slightly in each case but the arbitrary and nonuniform sampling designs appear no better than that of the uniform sampling case. The uniform sampling method required only $2 \times 17^3 = 9826$ floating point operations, while the nonuniform sampling method required on the order of $2 \times 9^4 \simeq 13,000$ floating point operations. The arbitrary sampling method required on the order of $1/64 \times 17^6 \simeq 380,000$ floating point operations. In looking back at the desired characteristics of this filter, the passsband and stopband regions comprise fairly equal areas of the (ω_1, ω_2) plane. Sampling on a uniform Cartesian grid allows an adequate number of samples in all bands and hence little seems to be gained through nonuniform spacing of samples.

2. Lowpass Filter No. 2

For the second example, another circularly symmetric 2–D lowpass filter characteristic is desired. In this case, the filter is to have unity gain in the passband $\{0 \leq \sqrt{\omega_1^2 + \omega_2^2} \leq 0.2\pi\}$ and zero gain in the stopband $\{0.4\pi \leq \sqrt{\omega_1^2 + \omega_2^2} \leq \pi\}$. This particular example was selected since it illustrates a case in which the desired passband encompasses a much smaller portion of the region K of the (ω_1, ω_2) plane than the stopband.
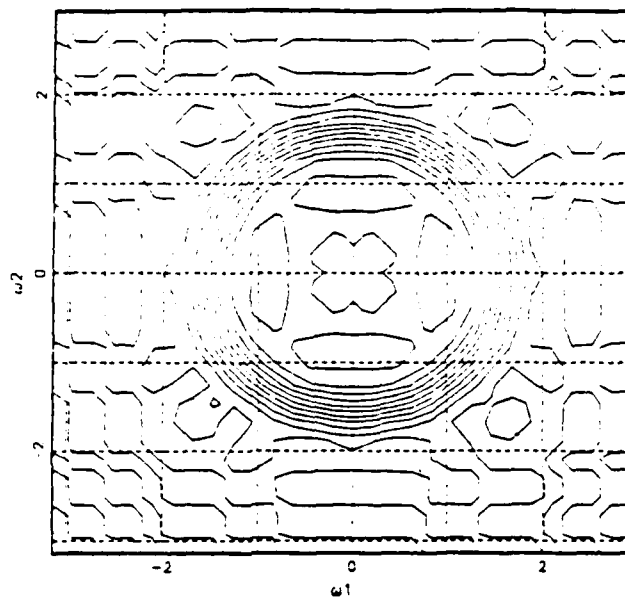
(a)
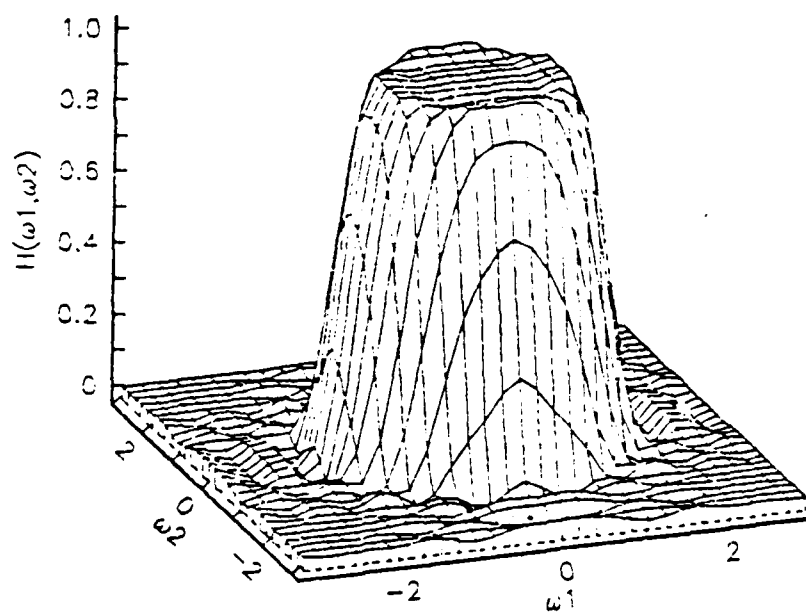
(b)

Figure 5.1. Frequency Response of Lowpass Filter Design No. 1,
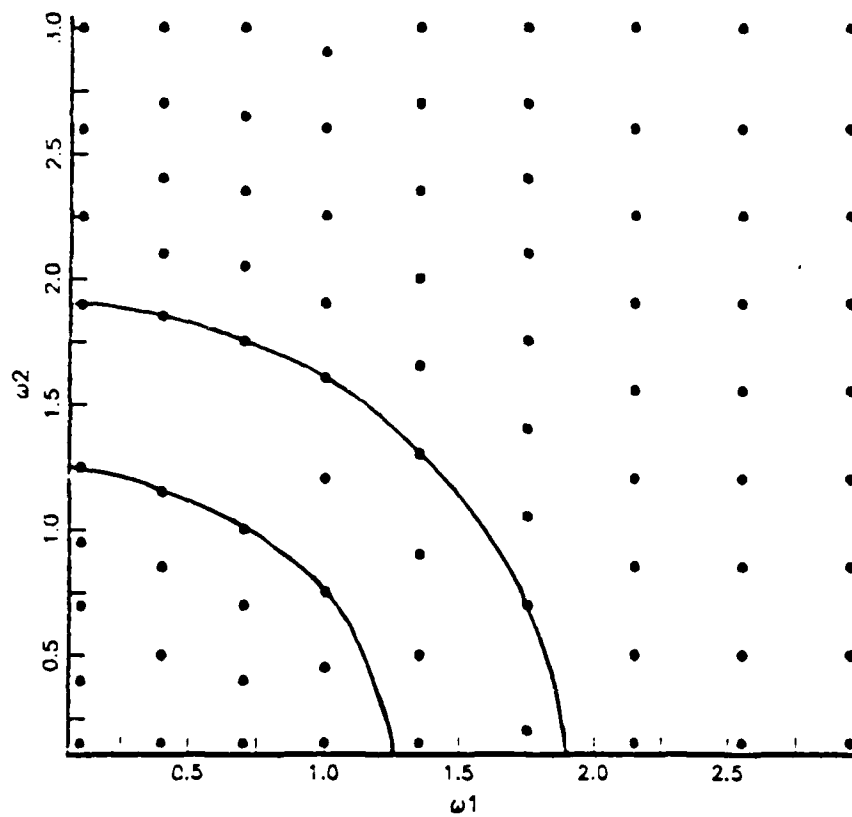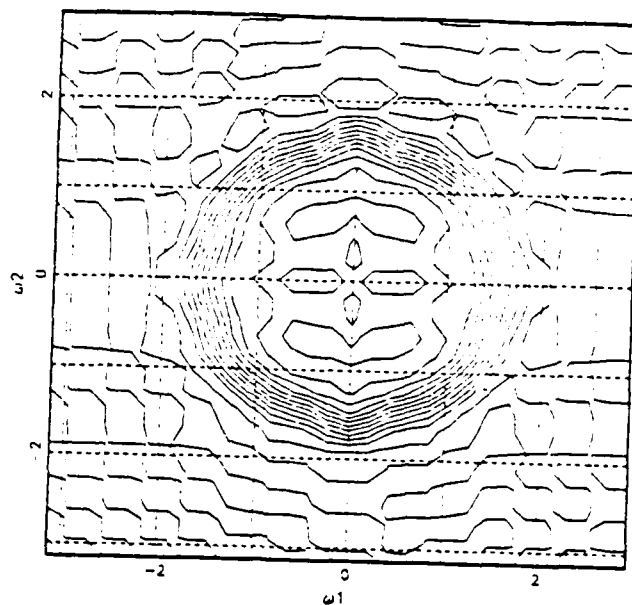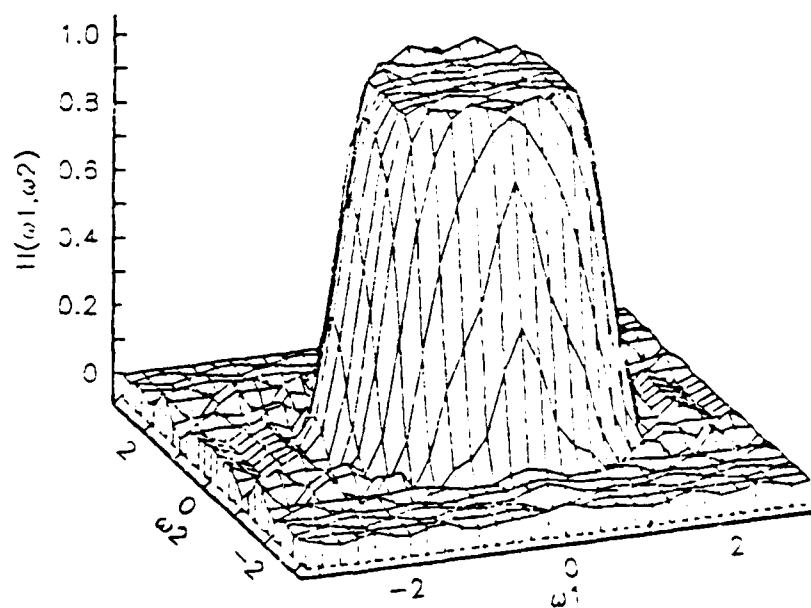Uniform Sampling Method.

51

Figure 5.2. Arrangement of Samples for Lowpass Filter
Design No. 1, Nonuniform Sampling Method.

52

(a)

(b)

Figure 5.3. Frequency Response of Lowpass Filter Design No. 1,
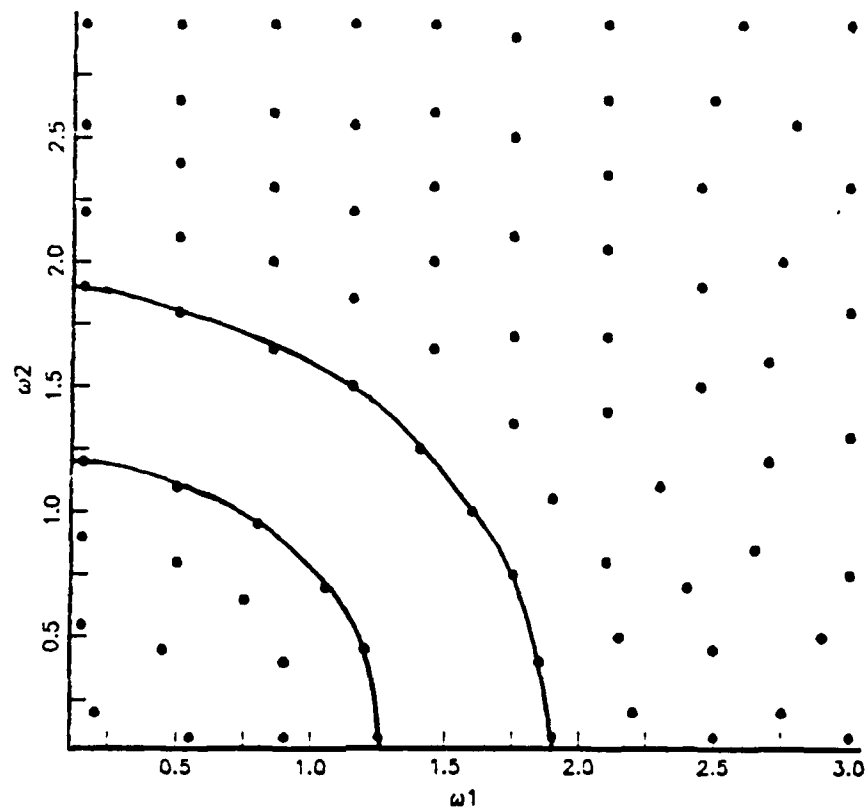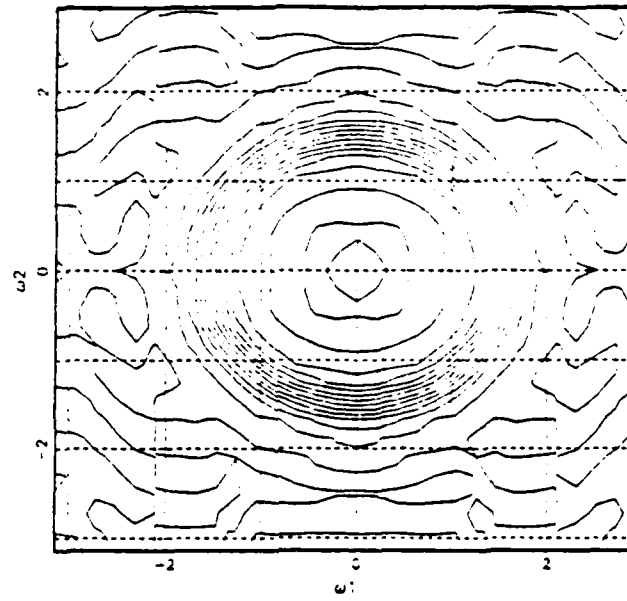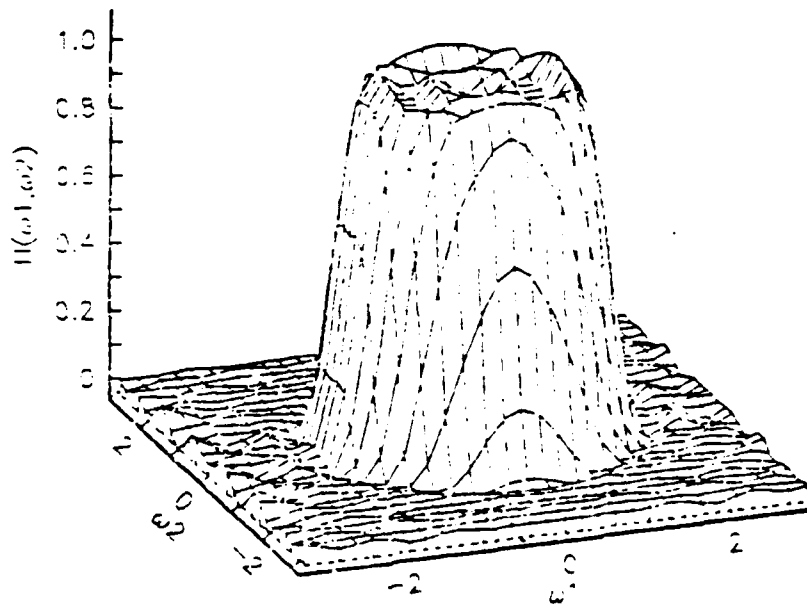Nonuniform Sampling Method.

53

Figure 5.4. Arrangement of Samples for Lowpass Filter
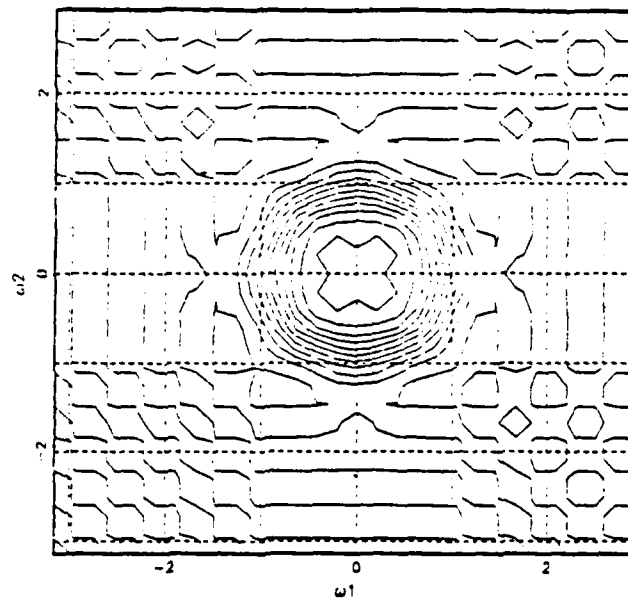Design No. 1, Arbitrary Sampling Method.

54

(a)

(b)

Figure 5.5. Frequency Response of Lowpass Filter Design No. 1,
Arbitrary Sampling Method.

55

The use of the uniform sampling method allows the placement of only four samples in the passband region. The resulting filter frequency response, shown in Figure 5.6, has a rather smooth passband and stopband characteristic, but the passband is smaller in area than desired and can be seen to have a rather squared shape in the contour plot.

The specified selection of frequency samples for the nonuniform sampling method is illustrated in Figure 5.7. Note that five samples are now placed within the desired passband region. Figure 5.8 represents the frequency response of the resulting filter. In this case, the size of the resulting passband more closely matches the desired characteristic and the passband is clearly more circular in shape. These characteristics are gained, however, at the expense of introducing a greater amount of ripple in the stopband.

Figure 5.9 shows a chosen arrangement of samples for the arbitrary sampling method. In this instance, six samples are now contained in the specified passband region while no samples are placed in the transition region. The resulting filter frequency response is illustrated in Figure 5.10.

Note that for the latter result, the passband is fairly flat and quite circular, and the roll-off is quite sharp. This occurs since the flexibility of the design allows frequency samples to be placed along the perimeters of the specified passband and stopband regions for the latter two design methods. Because of this, the two methods which allow flexibility in sampling can be used to provide greater control over the resultant width of the transition band. Of course, this is achieved at the expense of increased ripple. Since the arbitrary and nonuniform sampling methods do not require sampling on a Cartesian grid, they are generally better able to conform to non-rectangularly shaped contours such as circles.

(a)
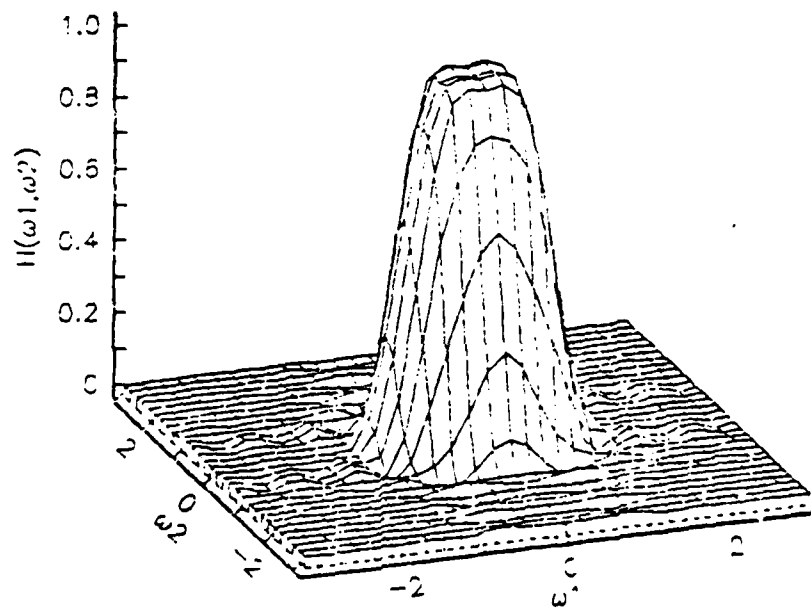
(b)

Figure 5.6. Frequency Response of Lowpass Filter Design No. 2,
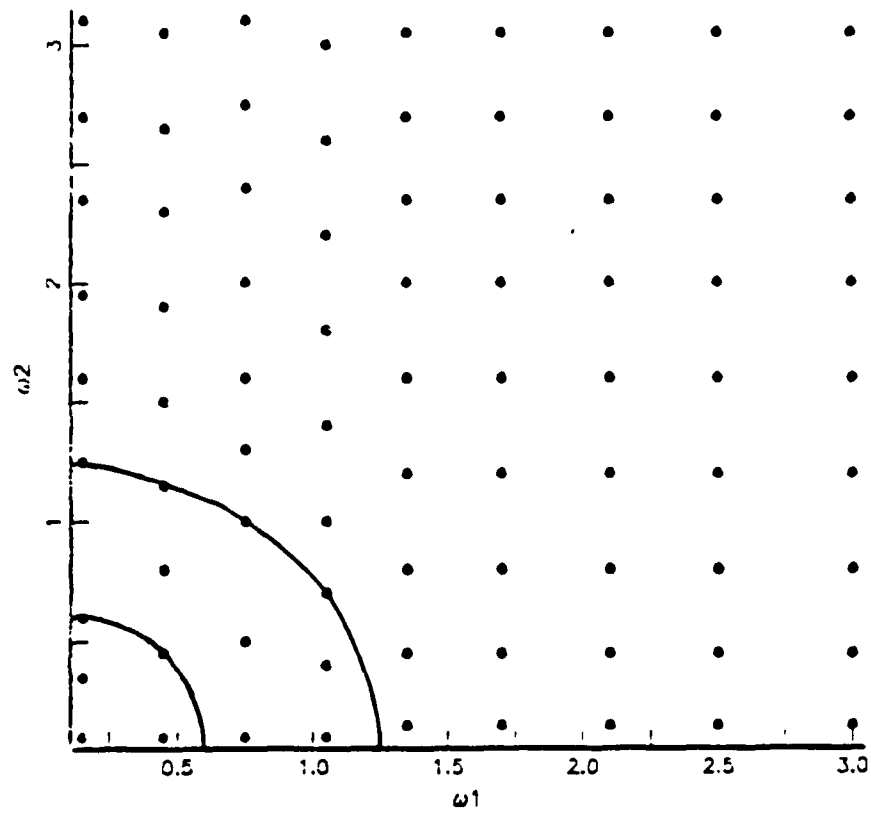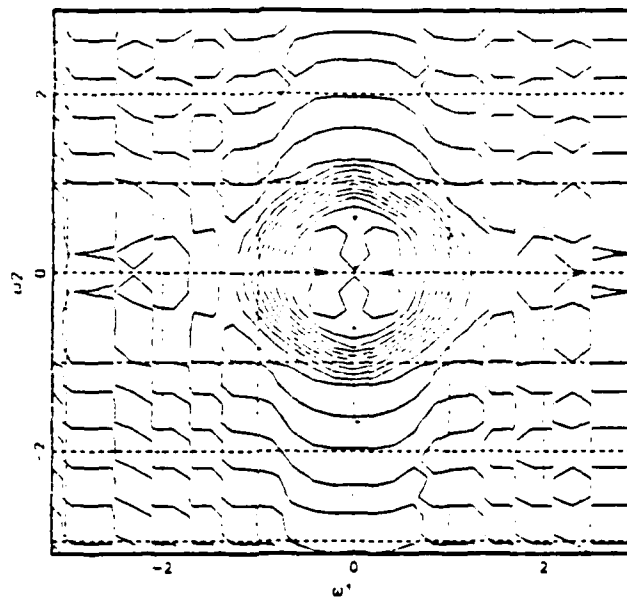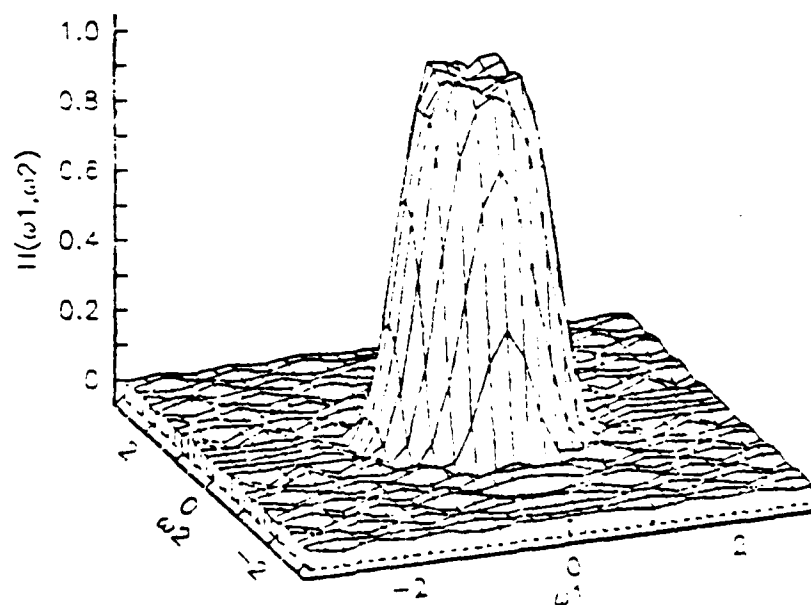Uniform Sampling Method.

57

Figure 5.7. Arrangement of Samples for Lowpass Filter
Design No. 2, Nonuniform Sampling Method.

58

(a)

(b)

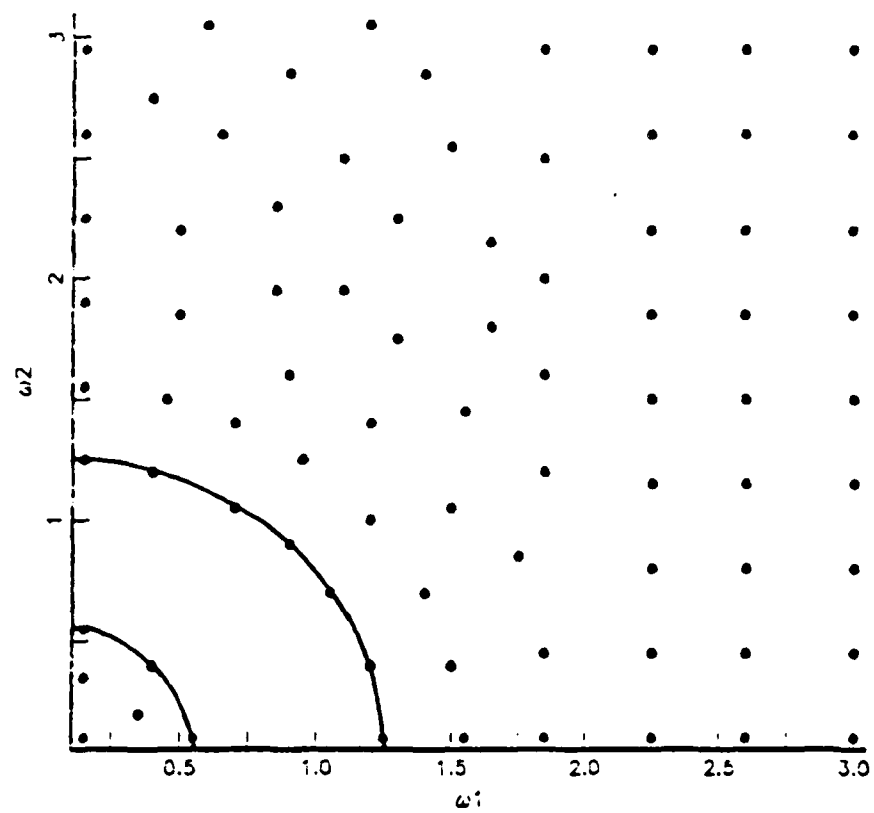Figure 5.8. Frequency Response of Lowpass Filter Design No. 2,
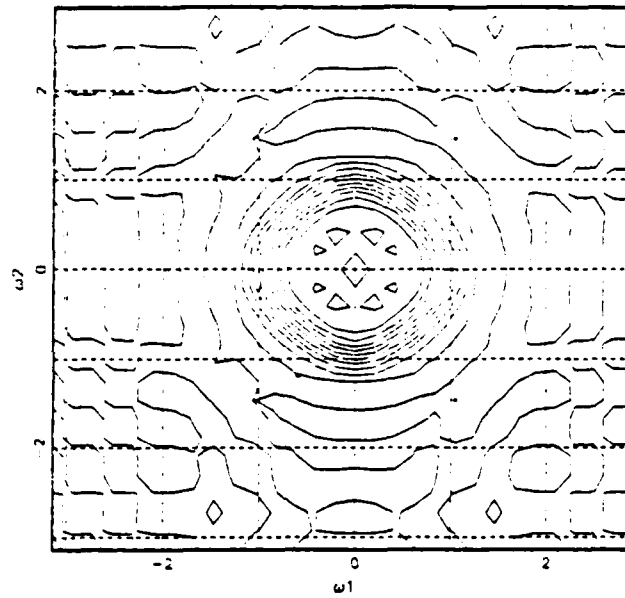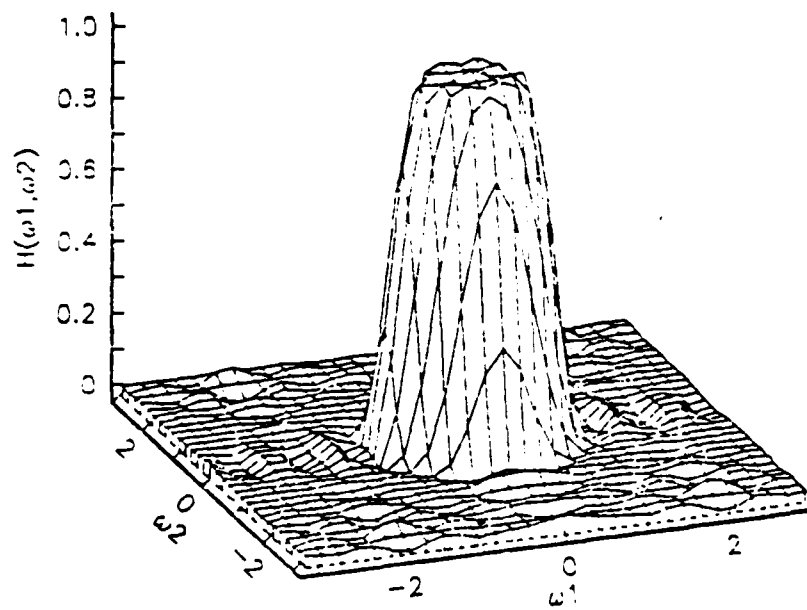Nonuniform Sampling Method.

59

Figure 5.9. Arrangement of Samples for Lowpass Filter
Design No. 2, Arbitrary Sampling Method.

60

(a)

(b)

Figure 5.10. Frequency Response of Lowpass Filter Design No. 2,
Arbitrary Sampling Method.

61

3. Bandpass Filter

The final design example illustrates how each of the frequency-sampling techniques can be used in the design of a 2-D circularly symmetric bandpass filter. Here, the desired filter has unity gain in the passband region $\{0.4\pi \leq \sqrt{\omega_1^2 + \omega_2^2} \leq 0.6\pi\}$, and zero gain in the stopbands defined by $\{0 \leq \sqrt{\omega_1^2 + \omega_2^2} \leq 0.2\pi\}$ and $\{0.8\pi \leq \sqrt{\omega_1^2 + \omega_2^2} \leq \pi\}$. The filter resulting from application of the uniform sampling method has the frequency response illustrated in Figure 5.11. For the nonuniform sampling method, the arrangement of samples shown in Figure 5.12 was used. The resulting filter frequency response is shown in Figure 5.13. Figure 5.14 illustrates the arrangement of samples selected for the arbitrary sampling method. The resulting filter has the frequency response of Figure 5.15.

In comparing these results, the nonuniform sampling method produces a somewhat flatter and wider passband than the uniform sampling method, but results in some notable ripples in the outer stopband which are not found in the uniform sampling case. The passband for the arbitrary sampling design is even more wide and flat, but at least for this particular selection of samples, has unacceptably high ripple in both stopband regions.

The point to be made in comparing these methods is not that the nonuniform or arbitrary sampling methods necessarily produce better filters, but rather that they possess more flexibility and hence can be used to enhance a particular desirable feature at the expense of another. The advantage gained through design flexibility may well outweigh the cost of additional computations for many design problems.

62
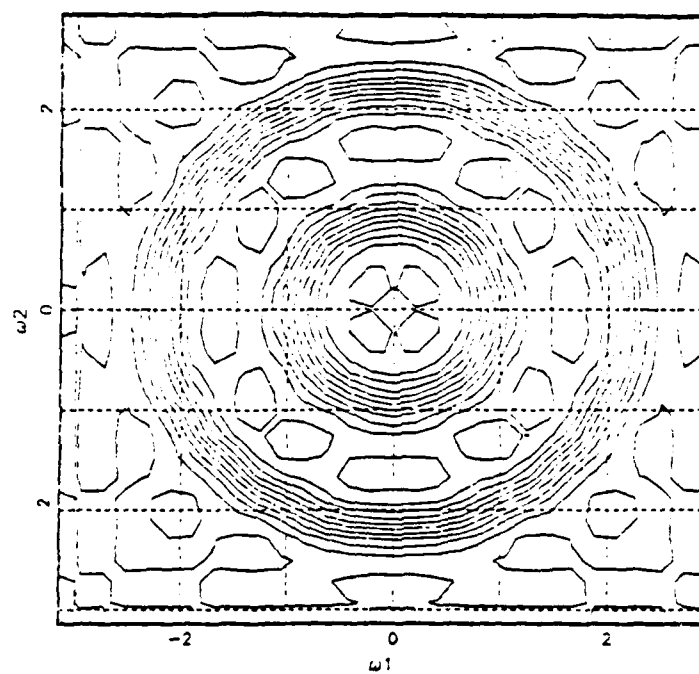
(a)
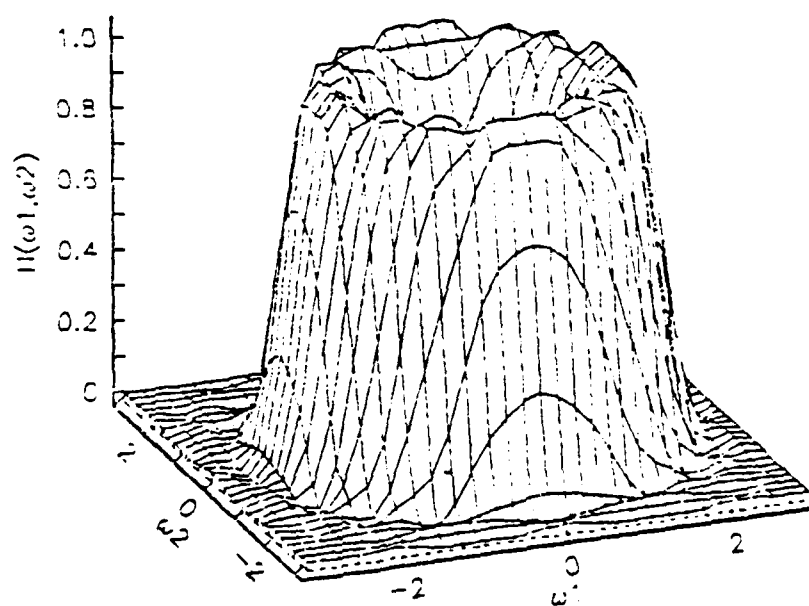
(b)

Figure 5.11. Frequency Response of Bandpass Filter Design,
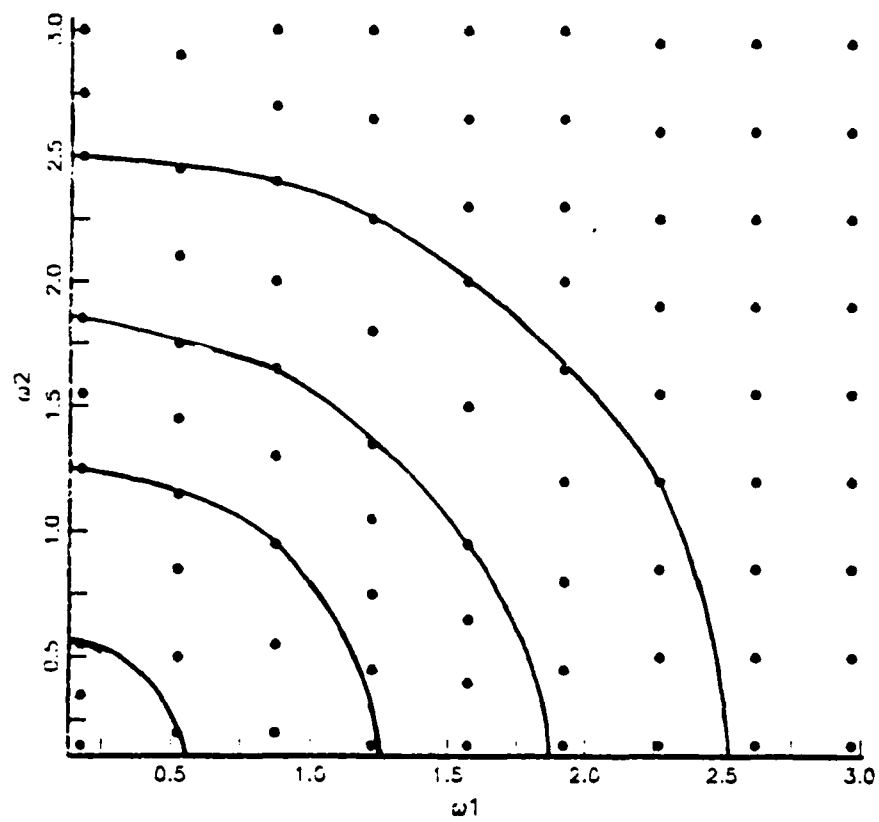Uniform Sampling Method.
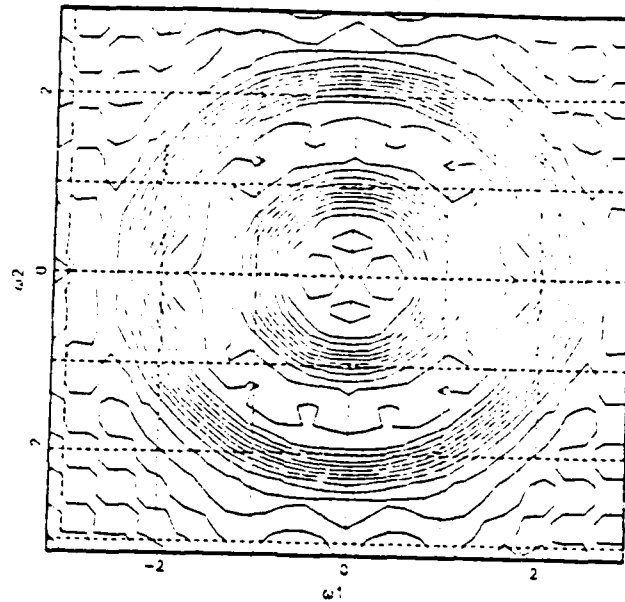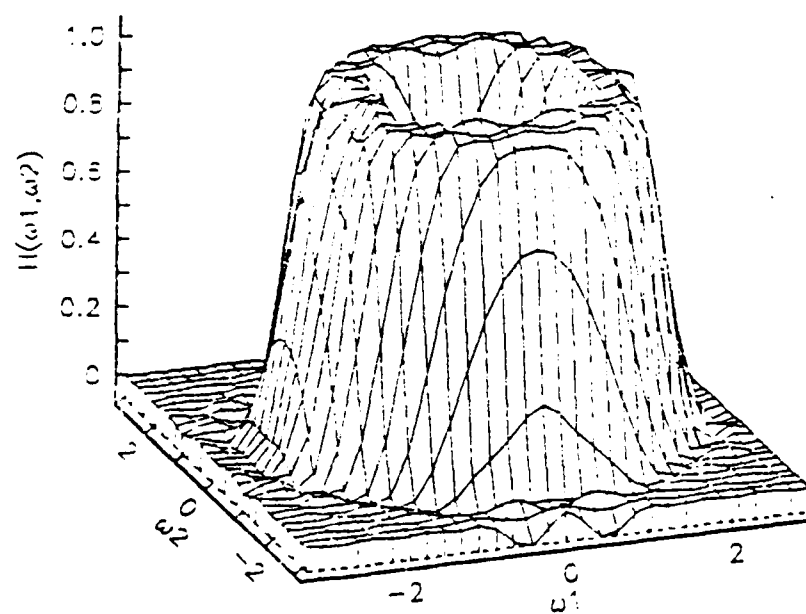
63

Figure 5.12. Arrangement of Samples for Bandpass Filter
Design, Nonuniform Sampling Method.

(a)

(b)

Figure 5.13. Frequency Response of Bandpass Filter Design,
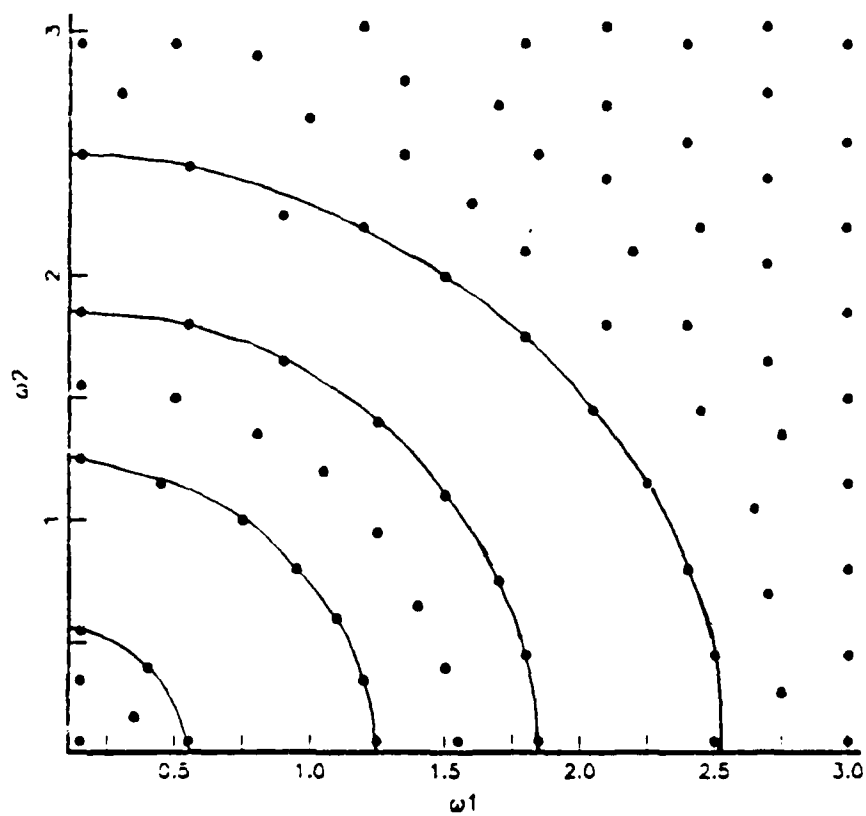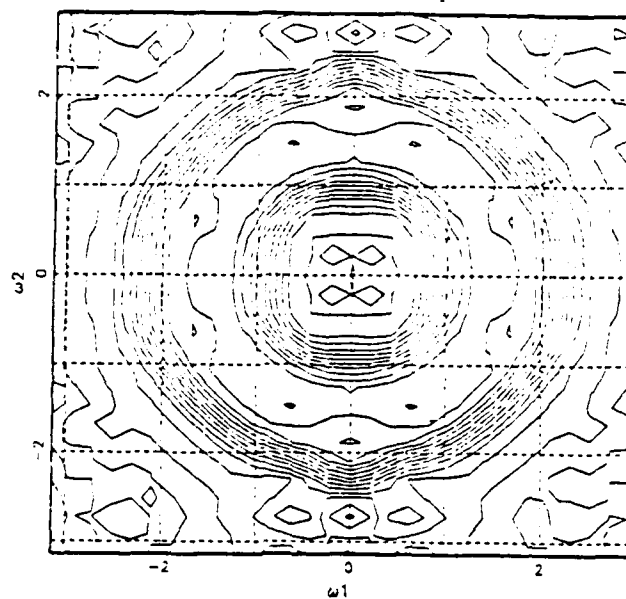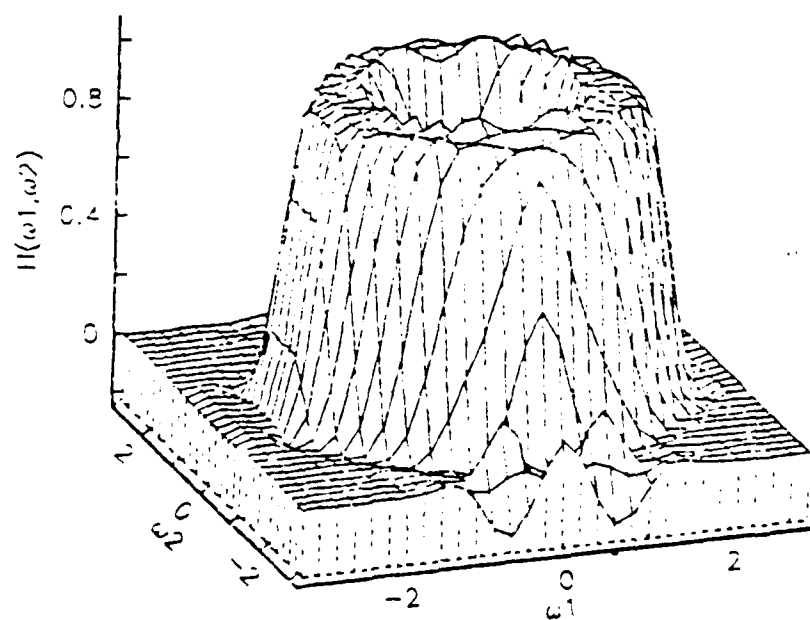Nonuniform Sampling Method.

65

Figure 5.14. Arrangement of Samples for Bandpass Filter
Design, Arbitrary Sampling Method.

(a)

(b)

Figure 5.15. Frequency Response of Bandpass Filter Design,
Arbitrary Sampling Method.

67

VI. CONCLUSIONS

In this thesis, various approaches to the frequency-sampling design of two-dimensional FIR digital filters were investigated. The traditional frequency-sampling approach, referred to as the uniform sampling method here, involves specifying the frequency response at sample points located at the vertices of a uniform Cartesian grid in the (ω_1, ω_2) plane. Through use of the IDFT, the method is quite computationally efficient, and can be made even more efficient if an FFT is used. The method possesses no flexibility in matching design parameters, however. Another frequency-sampling approach arises from specifying the frequency response at arbitrarily selected distinct samples in the (ω_1, ω_2) plane and then solving a large system of linear equations. This method, referred to as the arbitrary sampling method in this thesis, possesses great flexibility but is computationally intensive and prone to the occurrence of degeneracies in which a unique solution does not exist. Finally, a new frequency-sampling approach, termed the nonuniform sampling method, was introduced. This method somewhat constrains the location of frequency samples, but still provides some free parameters and is reasonably efficient in computation.

The results of this thesis reflect how computational efficiency and design flexibilty relate to the number of constraints on sample location. The methods described here which impose more constraints are more computationally efficient. Methods with less constraints were found to be more flexible in matching a desired frequency response.

A problem posed by both the arbitrary and nonuniform sampling methods is how to properly select the samples in order to achieve an acceptable filter. Various

68

algorithms for the design of 2-D optimal filters have been developed which involve the solution of linear equations, analogous to those for the arbitrary sampling method, in an iterative approximation process. Although they have been shown to converge, the algorithms necessarily take special measures to deal with degeneracies and are very computationally intensive. [Refs. 9, 10]

The nonuniform sampling method warrants a closer look because of its relative computational efficiency. Due to constraints on the sampling geometry, the nonuniform sampling method is probably not well-suited to the design of globally optimal filters. If an efficient algorithm can be developed which will locate a set of points satisfying the non-uniform method constraints, and if this set guarantees a filter which, while not necessarily optimal, nevertheless is of acceptable quality, the non-uniform sampling method developed in this thesis may prove to be of great significance.

APPENDIX

A 2–D FIR FILTER DESIGN BASED ON A BIVARIATE EXTENSION OF NEWTON'S POLYNOMIAL INTERPOLATION FORMULA

In the course of this thesis research, an extension of Newton's polynomial interpolation method in two variables was studied and applied to the design of two-dimensional digital filters. Such an extension of Newton's method was proposed by Diamessis [Ref. 11]. Its application as a 2–D FIR filter frequency-sampling design method was presented at the ICASSP 87 conference [Ref. 12] and is summarized here.

A. 2–D NEWTON INTERPOLATION

The bivariate Newton interpolation problem for the 2–D quadratic case can be described as follows. Given

1) a set of points in the (x, y) plane:

$$\mathbf{S}_f \triangleq \{(x_0, y_0), (x_0, y_1), (x_0, y_2), (x_1, y_0), (x_1, y_1), (x_2, y_0)\} \qquad (A-1)$$

2) a set of values at those points:

$$\mathbf{F} \triangleq \left\{ \begin{array}{lll} f(x_0, y_0) \triangleq f_{00}, & f(x_0, y_1) \triangleq f_{01}, & f(x_0, y_2) \triangleq f_{02}, \\ f(x_1, y_0) \triangleq f_{10}, & f(x_1, y_1) \triangleq f_{11}, & f(x_2, y_0) \triangleq f_{20} \end{array} \right\} \qquad (A-2)$$

3) a quadratic 2-D polynomial:

$$p(x, y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 \qquad (A-3)$$

find

$$\mathbf{a} \triangleq [a_{00}, a_{10}, a_{01}, a_{20}, a_{11}, a_{02}]^T \qquad (A-4)$$

70

such that the following interpolating conditions are satisfied

$$p(x_0, y_0) = f_{00}, \quad p(x_0, y_1) = f_{01}, \quad p(x_0, y_2) = f_{02} \left.\right\}$$
$$p(x_1, y_0) = f_{10}, \quad p(x_1, y_1) = f_{11}, \quad p(x_2, y_0) = f_{20} \left.\right\} .$$

$$(A - 5)$$

If the polynomial of Eq. (A-3) is rewritten in the form

$$p(x, y) = c_{00} + c_{10}(x - x_0) + c_{01}(y - y_0)$$

$$+ c_{20}(x - x_0)(x - x_1) + c_{11}(x - x_0)(y - y_0) \qquad (A - 6)$$

$$+ c_{02}(y - y_0)(y - y_1),$$

then the interpolation problem involves finding the coefficient vector

$$\mathbf{c} \triangleq [c_{00}, c_{10}, c_{01}, c_{20}, c_{11}, c_{02}]^T \qquad (A - 7)$$

instead of $\mathbf{a}$. The system of equations formed by application of Eq. (A-2) to Eq. (A-6) is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & (x_1 - x_0) & 0 & 0 & 0 & 0 \\ 1 & 0 & (y_1 - y_0) & 0 & 0 & 0 \\ 1 & (x_2 - x_0) & 0 & (x_2 - x_0)(x_2 - x_1) & 0 & 0 \\ 1 & (x_1 - x_0) & (y_1 - y_0) & 0 & (x_1 - x_0)(y_1 - y_0) & 0 \\ 1 & 0 & (y_2 - y_0) & 0 & 0 & (y_2 - y_0)(y_2 - y_1) \end{bmatrix} \begin{bmatrix} c_{00} \\ c_{10} \\ c_{01} \\ c_{20} \\ c_{11} \\ c_{02} \end{bmatrix}$$

$$= \begin{bmatrix} f_{00} \\ f_{10} \\ f_{01} \\ f_{20} \\ f_{11} \\ f_{02} \end{bmatrix}$$

$$(A - 8)$$

or, in vector-matrix notation,

$$\mathbf{N}_{2D} \mathbf{c} = \mathbf{f}. \qquad (A - 9)$$

71

Since N_{2D} is lower triangular, c can be solved for recursively. Also, N_{2D} is non-singular provided the x_i are distinct and the y_i are distinct. Hence a unique solution for c can always be found. The coefficient vector a is related to c by the relation

$$c = U_{2D} a, \qquad (A-10)$$

where U_{2D} is an upper triangular matrix which transforms the Newton coefficient vector c into the coefficient vector a. Therefore, a can be found through solution of the two triangular systems indicated by Eqs. (A-9) and (A-10).

B. NEWTON FILTER DESIGN METHOD

The selection of frequency samples for the Newton design method follows from the two variable Newton interpolation method discussed above. This approach is a 2–D extension of a one-dimensional filter design proposed by Schuessler [Ref. 13: p. 257]. If a square region of support on the first quadrant of extent $N \times N$ is initially specified for the filter impulse response, the frequency response of such a 2–D FIR filter with linear phase can be expressed as

$$H(\omega_1, \omega_2) = \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} h(n_1, n_2) e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}. \qquad (A-11)$$

A total of $(M+1)(M+2)/2$ frequency samples are then chosen on region K of the (ω_1, ω_2) plane, where $M = (N-1)/2$ and $K = \{(\omega_1, \omega_2), 0 \leq \omega_1 \leq \pi; 0 \leq \omega_2 \leq \pi\}$. The selected samples are of the form

$$\{((\omega_{1_{k_1}}, \omega_{2_{k_2}}), k_2 = 0, 1, \ldots, k_1), k_1 = 0, 1, \ldots, N\}. \qquad (A-12)$$

A support set of the above form is referred to as triangular. Note that the samples need not be uniformly spaced. Since there is no constraint such that $\omega_{i_0} < \omega_{i_1} < \omega_{i_2} \cdots$, the sampling geometry is not necessarily of a triangular shape on region

72

K. Two such support sets are shown in Figure A.1. Given a specified value for $H(\omega_{1_k}, \omega_{2_k})$ at each sample location, the above support set ensures existence of a unique solution to the design problem.

The Newton design imposes the quadrant symmetry condition

$$h(n_1, n_2) = h(N - 1 - n_1, n_2) = h(n_1, N - 1 - n_2). \qquad (A-13)$$

Application of Eq. (A-13) to Eq. (A-11) results in the expression

$$
\begin{aligned}
H_0(\omega_1, \omega_2) =& h(M, M) + 2 \sum_{n_2=0}^{M-1} h(M, n_2) T_{M-n_2}(\cos \omega_2) \\
&+ 2 \sum_{n_1=0}^{M-1} \sum_{n_2=0}^{2M} h(n_1, n_2) T_{M-n_1}(\cos \omega_1) T_{M-n_2}(\cos \omega_2),
\end{aligned}
\qquad (A-14)
$$

where $T_n(x)$ is the nth degree Chebyshev polynomial in x and $H_0(\omega_1, \omega_2)$ is the zero-phase frequency response related to $H(\omega_1, \omega_2)$ by

$$H(\omega_1, \omega_2) = e^{-j\omega_1 n_1} e^{-j\omega_2 n_2} H_0(\omega_1, \omega_2). \qquad (A-15)$$

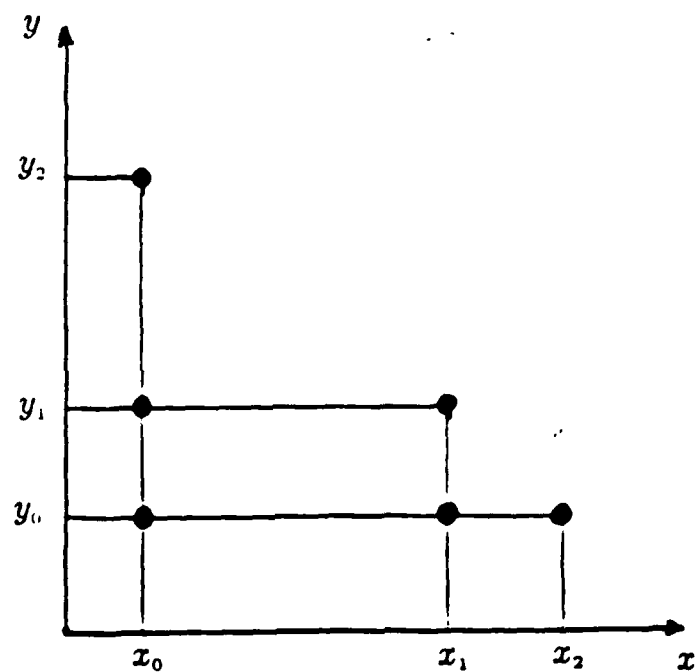Now, if the impulse response region of support is restricted such that

$$h(n_1, n_2) = 0, \quad |n_1 - M| + |n_2 - M| > M, \qquad (A-16)$$

and if the transformations

$$\cos \omega_1 = 1 - 2x \quad \text{and} \quad \cos \omega_2 = 1 - 2y \qquad (A-17)$$

are applied, then the zero-phase frequency response of Eq. (A-14) is mapped into a bivariate polynomial of the Newton form

$$
\begin{aligned}
\tilde{H}(x, y) =& c_{00} + \sum_{i=1}^{M} c_{i0} \prod_{p=0}^{i-1} (x - x_p) + \sum_{j=1}^{M} c_{0j} \prod_{q=0}^{j-1} (y - y_q) \\
&+ \sum_{i=2}^{M} \sum_{j=1}^{i-1} c_{i-j,j} \prod_{p=0}^{i-j-1} (x - x_p) \prod_{q=0}^{j-1} (y - y_q).
\end{aligned}
\qquad (A-18)
$$

73
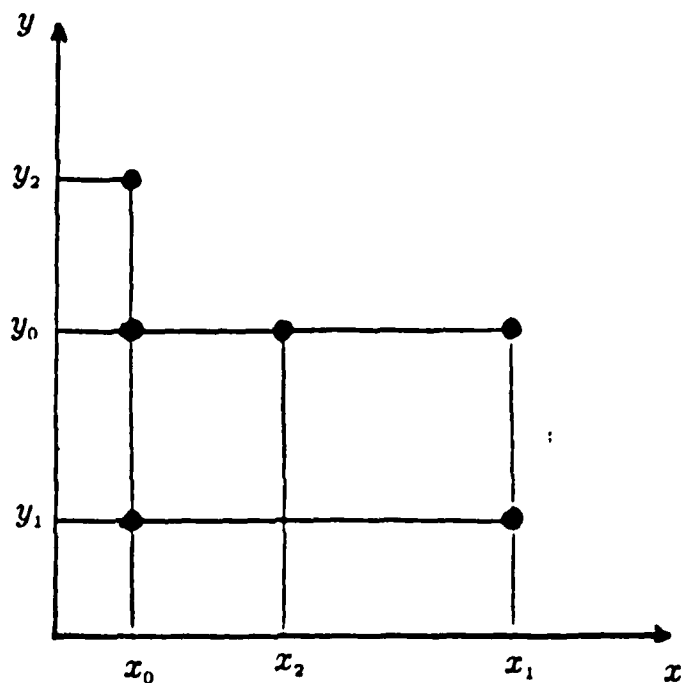
Figure A.1. Possible Sample Arrangements for Bivariate Newton
Interpolation, Quadratic Case.

74

The design problem involves first mapping the specified frequency samples $(\omega_{1_k}, \omega_{2_k})$ into samples (x_k, y_k) through the relations of Eq. (A-17), and then using the Newton interpolation method previously described to find the coefficients c_{ij}. By applying Eq. (A-17) to Eq. (A-18), and through the substitution

$$\cos\omega_i = \frac{e^{j\omega_i} + e^{-j\omega_i}}{2} \Rightarrow \frac{z_i + z_i^{-1}}{2},$$

a filter structure of the form

$$
\begin{aligned}
H(z_1, z_2) = &\, c_{00} z_1^{-M} z_2^{-M} + \sum_{i=1}^{M} c_{i0} z_1^{-(M-i)} z_2^{-M} \prod_{p=0}^{i-1} -\frac{1}{4}(1 - \mu_{1p} z_1^{-1} + z_1^{-2}) \\
&+ \sum_{j=1}^{M} c_{0j} z_1^{-M} z_2^{-(M-j)} \prod_{q=0}^{j-1} -\frac{1}{4}(1 - \mu_{2q} z_2^{-1} + z_2^{-2}) \\
&+ \sum_{i=2}^{M} \sum_{j=1}^{i-1} \Big\{ c_{i-j,j} z_1^{-(M-i+j)} z_2^{-(M-j)} \cdot \\
&\quad \prod_{p=0}^{i-j-1} -\frac{1}{4}(1 - \mu_{1p} z_1^{-1} + z_1^{-2}) \prod_{q=0}^{j-1} -\frac{1}{4}(1 - \mu_{2q} z_2^{-1} + z_2^{-2}) \Big\},
\end{aligned}
$$

$$(A-19)$$

results, where $\mu_{1p} = 2\cos\omega_{1p}$ and $\mu_{2q} = 2\cos\omega_{2q}$.

C. DISCUSSION

The advantages hoped to be gained by design of the Newton filter were computational efficiency through the recursive nature of the solution and use of the permanence property [Ref. 11]. The latter property allows the specification of additional sample values for an interpolation problem without having to redetermine previously computed coefficients. These advantages, however, apply only if the filter is implemented using the structure implied by Eq. (A-19), where the coefficients c_{ij} represent the gains. Such a structure, however, requires more delays than a direct form structure of the same order which uses the impulse response coefficients $h(n_1, n_2)$ for the filter gains, and hence is not very attractive.

75

An alternate procedure might be to use the Newton method to find the impulse response coefficients and implement the filter in direct form. Obtaining the impulse response coefficients of the filter from the set of coefficients c_{ij} requires solution of two triangular system analogous to those of Eqs. (A-9) and (A-10). Carrying out this two-step process, solving first for c and then for $h(n_1, n_2)$ to get a direct form implementation, is equivalent to solving a system of equations via LU-decomposition. This particular system is the system of linear equations resulting from the direct application of the interpolating conditions to Eq. (A-11), where the limits of the summation in Eq. (A-11) are altered to reflect the Newton filter region of support. This is the same approach used in the arbitrary sampling method of Chapter III, if the matrix inversion is performed through use of LU-decomposition, and therefore cannot be considered any more computationally efficient.

The Newton method may be less significant in terms of computational efficiency or practical filter implementation than it is for studying the formulation of the two-dimensional frequency-sampling design problem. In Chapter III, it was shown that existence and uniqueness of a solution to the interpolation problem for arbitrarily placed samples and a fixed impulse response region of support was not guaranteed for the two-dimensional case. With specification of a triangular support set and a polynomial of the form of Eq. (A-6), the Newton method provided a unique solution. If another, less constrained choice of samples is specified and if the form of the polynomial is altered to include basis functions such that the interpolating problem becomes a lower triangular system of equations akin to Eq. (A-8) with a non-zero main diagonal, then a unique interpolating polynomial of the new form will exist. Since the specified basis functions for the polynomial interpolation problem directly relate to the region of support for the impulse re-

76

sponse in the filter design problem, the Newton method can be used to illustrate a connection between the particular sampling geometry and an impulse response region of support which will ensure a unique solution.

LIST OF REFERENCES

1. Oppenheim, A.V. and Schaefer, R.W., *Digital Signal Processing*, Prentice-Hall, 1975.

2. Hu, J.V. and Rabiner, L.R., "Design Techniques for Two-Dimensional Digital Filters," *IEEE Trans. Audio Electroacoust.*, Vol. AU-20, No. 4, October 1972, pp. 249-256.

3. Lim, J.S., *Two-Dimensional Signal Processing and Image Processing*, 1988 (to be published).

4. Rabiner, L.R., Gold, B., and McGonegal, C.A., "An Approach to the Approximation Problem for Nonrecursive Digital Filters," *IEEE Trans. Audio Electroacoust.*, Vol. AU-18, No. 2, June 1970, pp. 83-106.

5. Dudgeon, D.E. and Mersereau, R.M., *Multidimensional Digital Signal Processing*, Prentice-Hall, 1984.

6. Elliot, D.F. and Rao, K.R., *Fast Transforms*, Academic Press, 1982.

7. Cheney, E.W., *Introduction to Approximation Theory*, McGraw-Hill, 1966.

8. Davis, P.J., *Interpolation and Approximation*, Blaisdell, 1963.

9. Harris, D.B. and Mersereau, R.M., "A Comparison of Algorithms for Minimax Design of Two-Dimensional Linear Phase FIR Digital Filters," *IEEE Trans. Acoust., Speech, Signal Process.*, Vol. ASSP-25, December 1977, pp. 492-500.

10. Kamp, Y. and Thiran, J.P., "Chebyshev Approximation for Two-Dimensional Nonrecursive Digital Filters," *IEEE Trans. Circuits and Systems*, Vol. CAS-22, No. 3, March 1975, pp. 208-218.

11. Diamessis, J.E. and Therrien, C.W., "Modelling Unequally Spaced 2-D Discrete Signals by Polynomials," *Proc. 20th Asilomar Conf. Signals, Systems, Computers*, November 1986.

12. Diamessis, J.E., Therrien, C.W., and Rozwod, W.J., "Design of 2-D FIR Filters with Nonuniform Frequency Samples," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, April 1987, pp. 1665-1668.

13. Schuessler, W., "On Structures for Nonrecursive Digital Filters," *Arch. Elek. Übertragung*, Vol. 26, June 1972, pp. 255-258.

INITIAL DISTRIBUTION LIST

No. Copies

1. Defense Technical Information Center   2
   Cameron Station
   Alexandria, Virginia 22304-6145

2. Library, Code 0142   2
   Naval Postgraduate School
   Monterey, California 93943-5002

3. Department Chairman, Code 62   1
   Department of Electrical and Computer Engineering
   Naval Postgraduate School
   Monterey, California 93943

4. Professor Charles W. Therrien, Code 62Ti   6
   Department of Electrical and Computer Engineering
   Naval Postgraduate School
   Monterey, California 93943

5. Professor Roberto Cristi, Code 62Cx   1
   Department of Electrical and Computer Engineering
   Naval Postgraduate School
   Monterey, California 93943

6. Professor Robert D. Strum, Code 62St   1
   Department of Electrical and Computer Engineering
   Naval Postgraduate School
   Monterey, California 93943

7. Professor Jae S. Lim   1
   Department of Electrical Engineering and Computer Science
   Room 36–653
   Massachusetts Institute of Technology
   Cambridge, Massachusetts 02139

8. Dr. Dan Dudgeon   1
   MIT Lincoln Laboratory, Group 21
   244 Wood St.
   Lexington, Massachusetts 02173

9. Commander   2
   Operational Test and Evaluation Force (Code 623)
   Attn: LT W.J. Rozwod
   Norfolk, Virginia 23511-6388

10. Director, Research Administration, Code 012   1
    Naval Postgraduate School
    Monterey, California 93943

11. Professor John E. Diamessis   1
    Department of Electrical Engineering
    National Technical University
    Patision St. # 42
    Athens 147, Greece

END

Feb.

1988

DTIC